Heritability analyses of musculoskeletal conditions and exercise-induced pulmonary haemorrhage in Thoroughbred racehorses

By

Claire Elizabeth Welsh BVMS (Hons) MSc (VetSci) AHEA MRCVS

A thesis submitted for the degree of Doctor of Philosophy

School of Veterinary Medicine

College of Medical, Veterinary and Life Sciences

University of Glasgow

January 2014

ABSTRACT

Musculoskeletal conditions and exercise-induced pulmonary haemorrhage are commonly diagnosed in Thoroughbred racehorses worldwide, and have serious consequences for racehorse welfare and the racing economy. Despite increasing interest in the study of genetic susceptibility to disease from the veterinary research community as a whole over past decades, the Thoroughbred has been largely ignored as a study group. The availability of software capable of complex genetic analyses using large, unbalanced pedigrees has made the study of genetic susceptibility to disease a realistic prospect for veterinary researchers. This study aimed to complete preliminary analyses of the genetics of a number of important musculoskeletal conditions, and of exercise-induced pulmonary haemorrhage, in two different Thoroughbred populations. Multivariable regression analyses were performed to identify important environmental risk factors for each condition in each population, and heritability analyses were conducted. Genetic correlations between disease conditions were also investigated. Fracture, tendon injury, suspensory ligament injury, osteoarthritis and EIPH/epistaxis were found to be heritable traits in the Hong Kong population. Distal limb fracture, SDFT injury and epistaxis were also found to be heritable in the UK Thoroughbred population. Most heritability estimates were small or moderate in magnitude. Selective breeding strategies that identify those animals with low genetic risk could play a part in future efforts to reduce the incidence of these conditions, in conjunction with favourable environmental manipulations based on research evidence. Due to low heritability, most of the conditions studied here would reduce in incidence slowly if selective breeding were implemented, thus strategic environmental manipulations would be warranted alongside such longer-term efforts to provide effective incidence reductions.

A number of conditions were found to be positively genetically correlated, suggesting that risk reduction through breeding could reduce the risk of multiple diseases concurrently. For example, fracture and osteoarthritis were found to be positively genetically correlated (0.85 – 0.89) in the Hong Kong racehorse population. However, using the Hong Kong Thoroughbred population dataset, EIPH/epistaxis and tendon injury were negatively genetically correlated, which suggests that reduction in genetic risk of one of these may lead to increased genetic risk of the other.

Measures of the durability and performance of racehorses were investigated to assess whether they were heritable traits in the UK and Hong Kong racehorse populations, and to assess their relationship to the disease conditions studied. Selection based on more holistic measures of horse health and longevity such as 'career length' could be a more attractive prospect for stakeholders, as this could forego the need to select for many different traits individually. Career length, number of starts over the career, and the level of earnings were all heritable traits in both populations. These holistic traits were found to have variable relationships with the disease conditions studied in each population. These analyses are the first to assess the genetic contribution to risk for many important diseases in the Thoroughbred. They provide a starting point from which further investigations into the applicability of genetic manipulations could yield realistic and achievable tools for racing stakeholders to use to 'improve' the breed in future.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

APS	Average proportion of alleles shared
BAL	Bronchoalveolar lavage
BHA	British Horseracing Authority
BMD	Bone mineral density
BV	Breeding value
CI	Confidence interval
CL	Career length
DNA	Deoxyribonucleic acid
EBV	Estimated breeding value
ECA	*Equus caballus*
EIPH	Exercise-induced pulmonary haemorrhage
GEBV	Genomic estimated breeding value
GS	Genomic selection
GWAS	Genome-wide association study
h^2	Heritability
HK	Hong Kong
HKJC	Hong Kong Jockey Club
LD	Linkage disequilibrium
LRT	Likelihood ratio test
MC3	Third metacarpal bone
MCP	Metacarpophalangeal joint
MC/T3	Third metacarpal or metatarsal bone
MCMC	Markov-Chain Monte Carlo
MSI	Musculoskeletal injury
MTP	Metatarsophalangeal joint
NH	National hunt
NHF	National hunt flat
NS	Number of career starts
OA	Osteoarthritis
OC	Osteochondrosis
OCD	Osteochondritis Dissecans
OVE	Official Veterinary Examination
PH	Per horse dataset (BHA data)
PSBF	Proximal sesamoid bone fracture
QTL	Quantitative trait locus
r	Repeatability
REML	Residual maximum likelihood
REP	Repeated measures dataset (BHA data)
SDFT	Superficial digital flexor tendon
s.e.	Standard error
SL	Suspensory ligament
SNP	Single nucleotide polymorphism
TB	Thoroughbred
TBA	Thoroughbred Breeding Association
UK	United Kingdom
USA	United States of America
VO	Veterinary official
VT	Veterinary technician
WG	Winnings group

ACKNOWLEDGEMENTS

My first note of gratitude has to go to the BBSRC and the University of Glasgow for offering me a place on my Doctoral Training Grant program. Without this opportunity, I would have found moving into research too daunting a prospect after having left vet school and working in mixed practice. This program was the best possible introduction to the diverse world of veterinary research, and I thoroughly enjoyed the whole process. I hope similar schemes continue to be offered in future so that academically-minded vets have a pathway into research from practice.

Thank you to the British Horseracing authority and the Hong Kong Jockey Club, and in particular Anthony Stirk and Doctor Kenneth Lam for the provision of the datasets used throughout this thesis, and for their help at various stages of the analysis process.

Many thanks must go to my great supervisors on this project, Doctor Timothy Parkin and Professor Dominic Mellor. Thank you both for your support, encouragement, and for providing the relaxed atmosphere in which I work best. The whole epidemiology group at the University of Glasgow were friendly and welcoming, which helped me feel less of an imposter than I may have done otherwise, so for that I'd like to thank them all.

I'd like to sincerely thank Doctor Sarah Blott and Doctor Tom Lewis for the considerable amount of time and effort they put into trying to help me understand the genetic concepts crucial to this project. Without their help this PhD would never have been possible. Also, thank you to the other staff and students of the Animal Health Trust, who made my trips there fun and busy.

Lastly, a huge thank you must go to my husband Alex and my family and friends, who have been ceaselessly patient with me when I've wanted to moan if things were going wrong, or chatter away incessantly about genetics when things were going well. I had a wonderful support network around me, which helped make the PhD process enjoyable and achievable.

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis and performed all the works presented here.

*signature

Claire E. Welsh

Date

1. CHAPTER I. INTRODUCTION

1.1 General introduction

Horseracing is one of the oldest and most popular spectator sports worldwide. The popularity of betting on horses has continued to increase in recent years despite widespread economic difficulties in the developed world. Serious injuries to racehorses are lamentably common in all types of racing, and fuel much debate in the media on the ethics of the sport. Injuries cost the racing industry a great deal through wastage of horses, veterinary bills, and importantly, through negative publicity. Many studies have been conducted in the past to better understand the epidemiology and risk factors for injuries, but few have led to widespread changes in horse management, and none have subsequently been proven to have reduced injury incidence. Past studies have often differed in case definition, statistical power, and the population of horses studied (amongst other aspects), meaning that clear, coherent advice for stakeholders on how to reduce injury incidence has not been forthcoming.

The Thoroughbred has been selectively bred for over 300 years to produce a superior athlete with a unique physiology. It is the breed most frequently used for racing worldwide, and is the focus of this study. Recent advances in equine genetics, such as mapping of the *Equus caballus* genome by the Broad Institute in 2007 (using the DNA of a Thoroughbred mare called Twilight) and development of the Illumina Equine SNP50 BeadChip, have fuelled a wealth of research into genetic diseases of the horse (Wade et al., 2009) . Much progress has been made in identifying the causal mutations of many monogenic diseases; indeed genetic tests are now commercially available for many of these conditions, for example Foal Immunodeficiency Syndrome (Bannasch, 2008; Brosnahan et al., 2010). However, as in human medicine, understanding multifactorial diseases, i.e. those for which multiple genes and the environment interplay to lead to disease, poses a more difficult challenge to equine geneticists. Dissection of the sources of variation in a population with respect to a disease enables a better understanding of its aetiology, and can focus efforts to reduce incidence. The availability of computing power and user-friendly software mean that the study of any disease should now include analysis of the possible contribution of genetic susceptibility to the overall disease prevalence.

Fracture and other musculoskeletal injuries are undoubtedly heavily influenced by environmental factors, but, far from precluding their genetic analysis, this indicates that a holistic approach to risk analysis is warranted, especially in instances where environmental

risk factors are impractical or impossible to manipulate, for example sex, year of racing and racing yard of origin (Boden et al., 2007; Lam et al., 2007a; Ramzan and Palmer, 2011). Research-led modification of risk factors such as the hardness or composition of the racing surface could potentially lead to swift reductions in risk of injury to all horses, and therefore such modifications are to be welcomed. However, some environmental influences are likely to be difficult to record accurately and also to model. Reduction of the risk conferred through the genes could enable a reduction in disease incidence even where environmental manipulation is unsuccessful or unfeasible, provided the condition in question is, at least in part, heritable. A thorough understanding of both the environmental and genetic influences of a trait and how these sources of variation work together in production of a disease phenotype is an important goal of veterinary research. What follows constitutes an important milestone on the path towards this aim for a number of significant veterinary conditions, in two geographically separate populations for the purpose of maximising global applicability of the results.

The success of the racing industry as a whole hinges on the ability of owners, trainers and racing employees to maintain the health and welfare of their charges, in order that they may perform well. As such, sound, reliable advice on the prevention of important veterinary diseases must be generated, disseminated and acted upon by stakeholders.

1.2 Horseracing and the Thoroughbred

1.2.1 Horseracing industry

1.2.1.1 United Kingdom horseracing

Horseracing in the United Kingdom (UK) has a long and established history, with the first recorded races taking place around 1174 under the reign of King Charles II. Currently, UK racing has the second largest capital expenditure of any sport (following football), and plays a vital role in the British economy (BHA, 2009).

The British Horseracing Authority (BHA) is the governing and regulatory body of horseracing in the UK (www.britishhorseracing.com). The remit of the BHA covers all areas of Thoroughbred racing, including personnel and course licensing, racing rules, drug testing and horse identification verification. All of the 60 racecourses in the UK must be

constructed and maintained to standards prescribed by the BHA, which include rules on course width, length, maximum number of runners, obstacle construction and placement, watering facilities, and ground surface quality. The BHA also oversees the veterinary care of all horses attending UK racecourses for the purposes of competing in registered races.

At least two course-employed Veterinary Technicians (VTs) must be present at every flat race fixture, and at least three at every jump race fixture. The VTs are responsible for much of the hands-on care of horses at the racetrack, and work closely with the BHA-employed Veterinary Official (VO). There must be one VO at every race fixture in the UK, and their principal responsibilities include liaising with course stewards to ensure adherence to BHA rules, drug testing and recording of veterinary diagnoses. The VTs and VOs work together to ensure that horses that are unfit to race are identified to the course stewards so that they may be prohibited from competing, and to deliver veterinary care to any horses suffering illness or injury before, during, or after a race (www.britishhorseracing.com/resources/media/publications_and_reports/BHAGIS_31.03.12.pdf).

Before every race fixture, the VO and stewards of the course populate a list of horses to be examined before racing, and the identity of all horses arriving at the course is verified by microchip scanning and/or passport checks. Horses that are noted include any that are aged 14 years or older in jump races, and 11 years or older in flat races, those that were reported lame at their last fixture, those that had not raced in the previous 12 months, or those that failed to complete four out of their last six races. If the VO detects any medical reason why these horses are unfit to race, the trainer will be notified and these horses will be removed from the race. All unexpectedly poor performances in races must be reported by the jockey or trainer to the VO with their explanation of why the performance was suboptimal. If a veterinary reason is suspected, the VO or VTs can examine the horse. All diagnoses and prognoses gathered from a race fixture are entered by the VO into a central database using a tick-box system, and an optional 60-character free text entry section.

Racing in the UK is either competed on flat ground or over jumps. Flat racing occurs on turf or all-weather synthetic surfaces. The hardness (‘going’) of turf tracks must be monitored and reported before each race meeting, and excessively hard or soft conditions can lead to cancellation of racing. All courses must have watering systems that are employed to avoid excessively hard ground. All-weather tracks must be made of approved

substances and an annual report on their condition and various performance criteria must be submitted to the BHA. All jump racing is on turf. Jump races are either hurdle or steeplechases, and jump courses must be wider than flat race courses to allow for lay-bys adjacent to obstacles. Hurdle races have at least eight hurdles over a two mile distance, with one additional hurdle for each subsequent quarter mile of track. Hurdles are smaller and more angled than steeplechase fences. Steeplechase races have at least twelve large, upright fences per two mile course, with an additional six fences per additional mile. Some of these fences can be preceded by an open ditch, and at least one is usually a water jump. The number of runners permitted in each race type ('field size') is set and enforced by the BHA. Hurdle and steeplechase races are collectively known as National Hunt (NH) races, which also incorporates National Hunt Flat (NHF) events. NHF races serve as an introduction to racing for horses intended for jump racing in future, but are themselves without obstacles, and usually take place on hurdles courses after the hurdles have been removed.

The BHA regularly publishes descriptive facts and figures about the British racing industry. The following facts are drawn from the BHA Fact Book 2011/12 (Akesson et al., 2008). In 2011, there were 14435 horses in training at 825 registered training yards. Of these horses, 8636 (59.8%) were in training for flat racing only, 4541 (31.5%) for jump racing only, and 670 (4.6%) were in training for dual purpose racing (plus 588, 4.1% in training as hunter chasers). Table 1-1 gives the number of horses in training for various racing disciplines in the UK in 2011. Table 1-2 gives the number of training yards that housed each group size of horses in training, in 2011.

Table 1-1. Data from the BHA Factbook 2011/2012. The number of horses of different ages in training for different racing disciplines in the UK in 2011.

	Horse age				
	2 years old	**3 years old**	**4 years old**	**5+ years old**	**Total**
Flat racing only	3043	2587	1264	1741	8636
Jump racing only	0	109	587	3845	4541
Dual purpose	0	62	169	439	670
All horses	3043	2758	2020	6025	13847

Table 1-2. Data from the BHA Factbook 2011/2012. The number of training yards in the UK that had between 1 and 10, or 11 and 15 etc. horses in training.

	Number of horses in training yard						
	1-10	**11-15**	**16-20**	**21-30**	**31-40**	**41+**	**Total**
Flat racing	250	47	37	34	18	60	446
Jump racing	537	41	18	15	12	17	640
Dual purpose	252	3	1	0	0	0	256

In 2011, 4635 live foals were produced by the UK's registered studs. UK racehorses are exported all over the world, and in 2011, of a total of 3633 exported Thoroughbreds, 1261 were sent to France, and 106 were sent to Hong Kong. The remaining 2266 horses were exported to more than 14 other countries on four continents.

The racing industry employed 4358 full time personnel and 2124 part time personnel in 2011. In that year, 896 flat racing fixtures were held, 570 jump racing fixtures and 3 mixed fixtures. In terms of racing surfaces, in 2011, 2245 races were run on all-weather tracks (20355 starts), 4025 on flat turf (39237 starts) and 3873 races were run over jumps (34735 starts).

The field size in UK racing in 2011 varied by race type as shown in Table 1-3.

Table 1-3. Data from the BHA Factbook 2011/2012. The number of races of different field sizes, of flat and jump racing in the UK in 2011.

	Field size				
	1-5	**6-10**	**11-15**	**16+**	**Total**
Flat racing	632	3346	2030	262	6270
Jump racing	673	1995	1022	187	3877

The BHA works alongside Wetherbys, which among many other duties, is responsible for maintenance of the General Stud Book in the UK (www.wetherbys.co.uk). The first edition of the stud book was published in 1791, and has been published every four years since. Admittance to the stud book is only allowed for those horses whose parents are already registered, and that have not been conceived using artificial insemination. Since 1999, all racehorses have been required to be microchipped to verify identity, and since 2001 DNA-based parent verification has been implemented.

The training, management and veterinary care of racehorses in the UK is predominantly outwith the control of the BHA, and horses are undoubtedly subjected to an almost limitless range of different environmental conditions. There is currently no provision for systematic recording of management techniques and training regimens of all registered racehorses in the UK.

1.2.1.2 Hong Kong horseracing

The Hong Kong Jockey Club (HKJC), established in 1884, became a professional organisation in 1971 (http://www.hkjc.com/home/english/index.asp). Unlike in the UK, the HKJC is responsible for all aspects of horseracing, including the housing, training and medical management of all Thoroughbreds in Hong Kong. The HKJC also licenses jockeys, trainers and trainers' assistants, and regulates which horse names are available for use. The sport continues to grow in popularity, generating HK$71,000 million in betting over the 2009/2010 season, and attracting in excess of 400,000 spectators in 2011 (HKJC, 2010, 2011). There are two racecourses in Hong Kong, both of which are flat. Happy Valley hosts most racing fixtures on its turf track, and Sha Tin is the site of most training, using both turf and all-weather tracks.

The HKJC annually houses around 1500 horses at Sha Tin racecourse in the New Territories. Around 400 horses are retired from racing annually, often to commercial riding establishments. Compulsory retirement is enforced by the HKJC for horses reaching 11 years old, those with a low performance rating at the end of a racing season, horses that are blind or missing one eye, following three episodes of epistaxis or three episodes of cardiac arrhythmia. The owner or trainer can opt to voluntarily retire a horse at any time. The racing year commences in July, with most races held between September and June. Approximately 80 race meetings and 700 races are held annually, with the majority run on turf.

The Veterinary Regulatory and Veterinary Clinical Departments conduct the veterinary care of all racehorses. Horses that are eight years old at the end of a racing season are required to pass a clinical examination before they are allowed to race in the following season. Any horse that has not raced for 12 months must undergo a barrier trial and pass a clinical examination before it can be permitted to race. One of the chief methods used by the HKJC for ensuring horse health is the Official Veterinary Examination (OVE) system.

OVEs can be issued by the chief stipendiary steward or official regulatory veterinary officer when a horse gives an unacceptable performance in a race, a veterinary problem is detected in a post-race examination, or when the regulatory department is advised by the clinical department of a horse with a significant diagnosis. The horse's trainer is informed of the OVE, and must work with the veterinarians to rectify the problem and restore the horse to health (where possible), before it can be resubmitted for an OVE, which it must pass before being allowed to resume racing. The trainer may alternatively decide to retire the horse. The OVE consists of an examination at gallop and a barrier trial, and a further examination one hour later. Horses with only minor problems may only be required to pass a clinical examination without a gallop test.

All episodes of epistaxis must be reported to the HKJC veterinary team on the same day that it occurred. Tracheobronchoscopy is not routinely performed to confirm exercise-induced pulmonary haemorrhage, and the horse is monitored for further episodes and banned from racing for a period of three months. A third episode of epistaxis requires compulsory retirement. Cardiac arrhythmias are diagnosed either by auscultation or electrocardiography. Following a first or second episode of arrhythmia, the horse is banned from racing for six months, and a third episode requires retirement.

1.2.2 Injuries sustained during racing and training

1.2.2.1 Overview

The investigation of the incidence of injuries to racehorses, and their associated risk factors, has frequently been conducted using rather broad case definitions in the past. Table 1-4 lists a number of these investigations with the case definitions used. Use of such broad definitions enables a general overview of the condition, and generates generic risk factors, but can be misleading when the reader is interested in a more specific injury, which may contribute few cases to the amalgamated group.

Table 1-4. Examples of published literature that have included case definitions pertaining to injuries.

Subject	Definition	Published Literature
Generic groups:		
Injury	Any injury or medical event diagnosed by observation and physical examination	Cohen et al. 2000, Pinchbeck et al. 2004, Stephen et al. 2003
Lameness	Lameness from all causes, based on owner 'diagnosis'	Murray et al. 2010
Days lost from training	Days lost from training due to any cause	Dyson et al. 2008, Taylor et al. 2012
Musculoskeletal Injury (MSI)	Trainer or veterinary diagnoses of injury or abnormality to any part of the musculoskeletal system	Cogger et al. 2008, Ramzan et al. 2011
Lower limb MSI	Injury to any part of a limb from the carpus/tarsus distally, which was associated with training and resulted in >7 day interruption to training	Perkins et al. 2005
Serious MSI	Veterinary-diagnosed MSI during or immediately following a race, resulting in euthanasia or failure to race for 6 months from the date of injury	Bailey et al. 1998
Catastrophic MSI	MSI necessitating euthanasia during or immediately following a race	Beisser et al. 2011
Fatality	Euthanasia or death at the racetrack, or within 24hrs of a race	Boden et al. 2007, Johnson et al. 1994
Fracture related groups:		
Fracture	Veterinary diagnosed fracture	Ely et al. 2004, Kaneko et al. 1996, Oikawa et al. 2005, Verheyen et al. 2007
Fatal distal limb fracture	Distal limb fractures (inclusive of carpus/tarsus) leading to euthanasia at the racetrack	Parkin et al. 2004
Third metacarpal/metatarsal (MC/T3) fracture	Fracture of any part(s) of the third metacarpal or metatarsal bones	Boyde et al. 2005, Jacklin et al. 2012, Parkin et al. 2006, Riggs et al. 1999, Stepnik et al. 2004, Whitton et al. 2010
Pelvic and tibial stress fracture	Veterinary diagnosed pelvic or tibial stress fractures	Verheyen et al. 2006
Proximal sesamoid bone fracture	Uni- or bilateral fracture of the proximal sesamoid bones	Anthenill et al. 2010, Kristoffersen et al. 2010, Kristoffersen et al. 2010
Tendon/ligament related groups:		
Tendon injury	Veterinary diagnosis of tendon injury leading to retirement	Lam et al. 2007, Lam et al. 2007
Tendonitis	Ultrasound (veterinary) diagnosis of tendon inflammation	Kalisiak et al. 2012
Suspensory apparatus (SA) injury	Injury to the suspensory ligament, proximal sesamoid bones of the forelimb, or distal sesamoidean ligaments, reported by trainer	Perkins et al. 2005
Suspensory ligament (SL) injury	Veterinary-diagnosed suspensory ligament injury	Ely et al. 2004, Kasashima et al. 2004
Superficial digital flexor tendon (SDFT) injury	Veterinary-diagnosed injury to the superficial digital flexor tendon	Ely et al. 2004, Kasashima et al. 2004, O'Meara et al. 2010, Perkins et al. 2005, Reardon et al. 2012, Thorpe et al. 2010

1.2.2.2 Fracture

Most injuries sustained by racehorses are injuries to the musculoskeletal tissues (MSI), and most MSI are fractures (Johnson et al., 1994; McKee, 1995; Stephen et al., 2003; Boden et al., 2005). Three main forms of fracture are recognised in horses, namely fatigue, monotonic and pathological forms (Riggs, 2002; Entwistle et al., 2008; Kristoffersen et al., 2010a). Pathologic fractures are those that occur either spontaneously or upon application of moderate forces to 'abnormal' diseased bone, for example following osteoporosis or infection. These are thought to represent a very small proportion of racecourse fractures, as serious pathology leading to weakening of bones would be likely to be addressed at an early stage, with the horse being withdrawn from racing. Monotonic fractures occur in previously healthy bone, and are the result of the application of excessive loads to bony tissues, such that the bone is deformed beyond its plastic threshold resulting in fracture (Riggs, 2002). Such injuries are often sustained following falls or collisions, and are identifiable by their unpredictable locations and fracture patterns, and their association with traumatic events. The third, and most common, form of fracture is the fatigue or 'stress' fracture. These have been recognised in human and equine medicine for many years, and are thought to be the result of accumulated damage to musculoskeletal tissues (Lanyon, 1987; Riggs, 2002; Shelton et al., 2003; Stepnik et al., 2004; Entwistle et al., 2008; Kristoffersen et al., 2010a; Ramzan and Palmer, 2011). Many studies have shown that training and racing activities allow adaptation of equine bone in particular locations, resulting in increased bone density (Nunamaker et al., 1990; Reilly et al., 1997; Stepnik et al., 2004; Boyde and Firth, 2005; Anthenill et al., 2010; Whitton et al., 2010; Beisser et al., 2011; Shi et al., 2011). This remodelled bone allows the formation of micro-cracks, which are protective against macroscopic fractures, provided they are not too numerous. With time, this microscopic damage will heal by osteoclastic resorption followed by osteoblastic bony proliferation, rendering the bone thicker in diameter than before, and less porous (Lanyon, 1987; Loitz and Zernicke, 1992; Riggs et al., 1999; Davidson, 2003; Boyde and Firth, 2005). Continued cyclical loading of bones suffering such microdamage, without sufficient rest periods to allow for healing, can lead to propagation of cracks and eventual catastrophic failure, i.e. a fatigue fracture. Due to their pathophysiology, these fractures are identifiable by their characteristic locations, evidence of histopathological changes of some duration in the fracture region, and the absence of a traumatic event in the history. Many studies have identified their presence in the past, but few have attempted to quantify the relative proportion of fatigue fractures to other forms of fracture in the racing

Thoroughbred (Nunamaker et al., 1990; Reilly et al., 1997; Riggs, 1999, 2002; Davidson, 2003; Parkin et al., 2006; Verheyen et al., 2006; Kristoffersen et al., 2010a; Ramzan and Powell, 2010; Whitton et al., 2010). Fractures of the third metacarpal or metatarsal bones were thought to be fatigue-related in the past, however more recent studies have suggested that only those fractures originating within the condylar groove show stress-related pathology, with a significant proportion of fractures originating outside the condylar groove (Riggs, 1999; Shelton et al., 2003; Reardon et al., 2012). Similarly, fractures of the proximal sesamoid bones have been shown to be associated with increased bone density in one study, but were not associated with an increased level of microdamage compared to unfractured sesamoid bones in an earlier study (Kristoffersen et al., 2010a; Shi et al., 2011).

Previous studies have reported fracture rates of 1.1 to 1.5 per 100 horse months in training, and more variable rates in racing, depending on the type of race (Ely et al., 2004; Dyson et al., 2008; Ely et al., 2009). The number of fractures per 1000 race starts has been reported as 0.4 in flat racing, rising to 9.2 in National Hunt racing (Stephen et al., 2003; Parkin et al., 2004a; Ely et al., 2009). These and other studies have shown an increased risk of fracture in jump racing compared with racing on the flat. Most fractures are serious, requiring euthanasia, retirement or a protracted period of rest (Johnson et al., 1994; Pinchbeck, 2004; Beisser et al., 2011).

Most studies of fractures in Thoroughbreds have identified that the distal forelimb is most at risk, however the predominating fracture types seen vary in different jurisdictions, on different surfaces and whether fracture occurred during racing or training. In the UK, one study found that overall, the lateral condyle of the third metacarpal bone was most commonly fractured, whereas on all-weather tracks and turf tracks, the proximal sesamoid bones and proximal phalanx were at greater risk, respectively (Parkin et al., 2004b). When fractures occur in training in the UK, pelvic and tibial fractures tend to predominate (Ely et al., 2004; Verheyen et al., 2006).

Many traditional epidemiological studies have been conducted in the past in the hope of identifying important modifiable horse-, course- or race-level risk factors for fracture. Course-level risk factors previously identified include the type of ground on which racing occurs, and the hardness of the ground (Parkin et al., 2004b; Oikawa and Kusunose, 2005; Kristoffersen et al., 2010b). One study suggested that the association between fracture risk

and ground hardness could be explained by race speed as an intervening variable, with higher speed races occurring on harder ground (Oikawa and Kusunose, 2005). Race type has also been associated with the risk of fracture or injury. Races with obstacles (i.e. hurdle and steeplechase) pose a greater risk than flat races (Bailey et al., 1998; Parkin et al., 2004b). Exercise history has been shown to be associated with MSI risk, and with fracture risk in training, however, the precise relationship has proven difficult to quantify (Cogger et al., 2006; Verheyen et al., 2006). One study found that fracture risk was reduced if the horse in question was the first parity of its mare, or if the mare was young when the horse was born (Verheyen et al., 2007). With the exception of hardness of the racing surface, these risk factors are difficult or impractical to modify.

1.2.2.3 Musculoskeletal injury

Musculoskeletal injuries (MSI) on the whole are the most common injuries of racing Thoroughbreds. For the purposes of this study, MSI are defined as musculoskeletal injuries not including fractures, which will be dealt with separately. Excluding fracture, tendon and ligament injuries are the most common MSI. Prevalence of MSI varies with race type, with the highest prevalence occurring where horses must clear large obstacles (i.e. steeplechase) (Ely et al., 2004; Pinchbeck, 2004). In flat racing, Bailey et al. (1998) reported a prevalence of 0.29% MSI, which rose to 2.91% in steeplechase racing. Stephen et al. (2003) reported a similar prevalence of 0.6% in flat racing and 3.1% in steeplechase racing.

The most common tendon to be injured by racehorses is the superficial digital flexor tendon (SDFT) (Thorpe et al., 2010). The SDFT plays a major role in ambulation through aiding flexion of the limb, and storing large amounts of energy which can be released in the swing phase of the gait. Injuries to this structure often occur following accumulation of damage to the tendon parenchyma, although the pathological mechanisms underlying this damage are poorly described (Thorpe et al., 2010). The prevalence of SDFT injuries is high, and re-injury is common (O'Meara et al., 2010). Of all retirements from racing in Hong Kong, 3.2% were caused by tendon injuries (Lam et al., 2007a). In 2011, Ramzan et al. found that 10.8% of injuries in racehorses at three training yards in Newmarket, UK, were SDFT injuries (Ramzan and Palmer, 2011). Many risk factors for SDFT injury have been identified in past studies. Both male sex and increasing age have been repeatedly found to be associated with increased SDFT injury risk (Kasashima et al., 2004; Perkins et

al., 2005a). Various aspects of a horse's training and exercise history have also been associated with SDFT injury risk; however, these relationships are complicated and have been contradictory in different studies (Perkins et al., 2005a; Cogger et al., 2006; Lam et al., 2007b; Reardon et al., 2012). Two studies found a significant association between the season in which racing occurred and the risk of SDFT injury (Perkins et al., 2005a; Reardon et al., 2012). In the UK study, racing in summer months was associated with increased risk; however, the New Zealand study found that SDFT injury risk was less in the period November to April compared with August to October, i.e. the risk was greater in winter versus summer. The reason for this discrepancy is unknown. Reardon et al. (2012) reported a number of significant race-level risk factors associated with increasing risk of SDFT injury, including increased race distance, time of the race (afternoon versus morning or evening), fewer runners per race, and carrying more weight (Reardon et al., 2012). It is impractical to consider modifying many of the associated risk factors found in these previous studies, such as gender and racing season. Many of the race and exercise history-related risk factors identified by Reardon et al. (2012) and others could be modified; however, this study was limited to Thoroughbreds racing over hurdles in the UK, thus the conclusions may not be applicable to other groups.

The suspensory ligament (SL) is the most common ligament to be injured in the racehorse. In the UK, 9.1% of MSI have been shown to be injuries to the SL (Ramzan and Palmer, 2011), and in flat racing in Japan 3.6% of injuries involved the SL (Kasashima et al., 2004). The incidence rate of SL injury in one study was 0.2-0.57 per 100 horse months in three and two-year olds respectively (Dyson et al., 2008). Injuries to the SL are common, but many studies have found them to occur at a lower frequency than injuries to the SDFT (Ely et al., 2004; Kasashima et al., 2004; Cogger et al., 2008; Dyson et al., 2008; Ramzan and Palmer, 2011). Although tendon and ligament injuries are often considered together as a single condition, the gross and histological anatomy, role and properties of these structures vary considerably, and there is evidence that risk factors for them differ, thus they should be considered as separate entities (Ely et al., 2004; Perkins et al., 2005a). There is a paucity of published literature examining the pathophysiology of damage to the SL specifically, or identifying risk factors for SL injury.

Osteoarthritis (OA) is a common diagnosis in racehorses, and describes degenerative changes to joints including articular cartilage erosion, subchondral bone sclerosis and osteophytosis (Brommer et al., 2004; Neundorf et al., 2010). The initial insult causing

joint inflammation could be trauma, osteochondrosis, age-related wear, or idiopathic, but the subsequent pathophysiological processes are similar irrespective of the primary cause. The metacarpophalangeal joint (MCP) is the most common site of OA in the horse, where histopathological changes can involve the articular surfaces of the third metacarpal bones, the proximal sesamoid bones or the proximal phalanx, or any combination of these (Neundorf et al., 2010). It is thought that this joint is more at risk than the metatarsophalangeal (MTP) joint due to its different biokinetics and loading patterns during exercise (Brommer et al., 2004). The prevalence of OA in the MCP joints of Thoroughbreds that died or were subjected to euthanasia within 60 days of racing was reported to be 33% in two and three year old horses, and increased with increasing age up to 6 years old (Neundorf et al., 2010). Literature citing the prevalence of OA in Thoroughbreds is sparse, as are studies of the risk factors of this condition. Both articular fractures and Osteochondrosis Dissecans (OCD), a developmental orthopaedic condition, are known to increase a horse's risk of developing OA. The goal in treatment of OA is to control pain, as the changes to the cartilage and subchondral bone are typically irreversible. To this end, many different pharmaceutical preparations are employed in OA treatment.

1.2.2.4 Exercise-induced pulmonary haemorrhage and epistaxis

Exercise-induced pulmonary haemorrhage (EIPH) is a very common condition of racehorses, with previous studies reporting prevalence to be up to 100% (Meyer et al., 1998; Davidson et al., 2011). EIPH is thought to be associated with poor performance, and has been linked to sudden death (Boden et al., 2005; Hinchcliff et al., 2005; Newton et al., 2005; Davidson et al., 2011). There are two competing theories surrounding the mechanisms by which haemorrhage is induced in the lung parenchyma. One theory suggests that forces generated by forelimb footfall during canter and gallop are transmitted as pressure waves to the dorso-caudal lung fields, which are recognised as the most common sites of haemorrhage (Schroter et al., 1998). An alternative theory, which has received greater attention in recent years, involves the generation of excessively high capillary transmural pressures as a result of very high cardiac output during exercise, leading to stress failure of capillary walls and resulting in haemorrhage (West and Mathieu-Costello, 1994). A further study implicated an impaired innate immune response of macrophages in EIPH-sufferers as a contributory feature of the condition (Michelotto et al., 2011). Diagnosis of EIPH relies upon the identification of sufficient numbers of haemosiderophages in broncheoalveolar lavage (BAL) fluid samples, but the invasive

nature of this procedure means that studies of EIPH often rely upon visual examination for bilateral epistaxis (blood at the nostrils) as a proxy, assuming that all bilateral epistaxis occurs as a result of EIPH (Stephen et al., 2003; Weideman et al., 2004). A recent study, however, has suggested that assessment of circulating levels of the hormone Angiotensin Converting Enzyme (ACE) may be of benefit in the diagnosis of EIPH without the need for BAL (Costa et al., 2012). Epistaxis following exercise is much less prevalent than EIPH, and is thought to occur only in serious EIPH cases (Takahashi et al., 2001; Williams et al., 2001; Hinchcliff et al., 2005; Newton et al., 2005). Williams et al. (2001) reported that the prevalence of epistaxis varied with different race types from 0.29 episodes per 1000 starts in National Hunt flat racing, up to 2.66 episodes per 1000 starts in steeplechase racing. Overall, this study reported that 7.8% of racehorses experienced at least one episode of epistaxis between 1996 and 1998. It may take a variable length of time for haemorrhage of pulmonary origin to reach the nostrils, thus studies in which epistaxis prevalence was considerably lower may have examined horses too soon following exercise and missed eventual cases of epistaxis (Stephen et al., 2003). Due to negative public perception of racehorses demonstrating epistaxis, and due to the perceived welfare implications of EIPH, drug therapy has been implemented in some areas and research conducted into the aetiology and prevention of the condition. Some racing jurisdictions, most notably the USA, allow the use of furosemide in racehorses to reduce the incidence of epistaxis at the racecourse. Furosemide is used in up to 92% of US Thoroughbreds, and its efficacy at reducing the incidence and severity of EIPH have been proven (Hinchcliff et al., 2009). This drug is banned in racehorses in the UK and Hong Kong due to its perceived performance enhancing effects.

Past risk factors that have been associated with increased risk of EIPH and/or epistaxis include colder air temperature, jump racing compared with flat racing, racing over hard ground, longer career length, being older than two years at the time of racing, short race length and being female (Lapointe et al., 1994; Takahashi et al., 2001; Newton et al., 2005; Hinchcliff et al., 2010).

1.3 Equine genetics

1.3.1 The Thoroughbred

The origins of the Thoroughbred date back nearly 400 years, when native British mares were crossed to imported Barbs, Turks and Arabians to produce a horse capable of great speed. The Thoroughbred Studbook was first published in 1791 (Weatherby, 1791), and since then the breed has remained effectively closed (Cunningham et al., 2001). Almost all paternal genetic diversity in the modern Thoroughbred can be traced back to one of three stud horses that were popular around 1700; the Godolphin Arabian, the Byerly Turk and the Darley Arabian. In modern times, nearly 95% of paternal lineages can be traced back to the Darley Arabian, through his great-great-grandson Eclipse (Cunningham et al., 2001). Only 28 mares contributed about 89% of modern Thoroughbred maternal lineages. The modern studbook contains in excess of 300,000 horses. Verification of the identity of individuals in the studbook was previously performed using blood-typing, but this practice was superseded by DNA profiling in 2001. Breeding of Thoroughbreds in the UK takes place principally at 300 full-time studs, with a further 4400 part-time breeding establishments also contributing to the population annually (BHA, 2009). Around 4000 foals are produced in the UK every year following natural mating, as artificial insemination is prohibited.

In a closed population such as that of the Thoroughbred, genetic diversity can suffer due to a lack of introgression of novel genes, especially where mating is non-random. The level of inbreeding in the Thoroughbred has been questioned in the past, with the suspicion that the mating of close relatives has led to suboptimal fertility, prevalent musculoskeletal disorders and a failure of performance to improve over generations (Corbin et al., 2010; Binns et al., 2012). To this end, a number of studies have estimated the current degree and rate of inbreeding, and the genetic diversity of the breed. Binns et al. (2012) found that the average inbreeding coefficient of the population had tended to increase with subsequent cohorts between 1961 and 2006, and also reported an increase in homozygosity, an indicator for inbreeding. Most of the increase in inbreeding over that time period had been concentrated in the last 10 years, which the authors described as not excessive, but worrisome. Prior to this study, Mahon et al. (1982) reported the average inbreeding coefficient over the 21.5 generations since the foundation of the breed to be 12.5%, which is the equivalent of mating grandfathers to granddaughters (Mahon and Cunningham,

1982). Despite this, no statistically significant association was found between inbreeding coefficient and fertility. Effective population size is the theoretical number of breeding individuals that would be required to produce the rate of inbreeding seen in the current population. A small effective population size is an indicator of a lack of genetic diversity. Corbin et al. (2010) estimated this number in the Thoroughbred to be around 180, having reached a peak two generations previously of 190 (Corbin et al., 2010). Although small compared with other horse breeds, the effective population size of the Thoroughbred was reported to have increased over the preceding 10 generations (Hamann and Distl, 2008). Other measures of genetic diversity include the number of alleles per locus, the average proportion of alleles shared by all individuals (APS), and the level of observed compared with expected heterozygosity. Cunningham et al. (2001) investigated all of these in the Thoroughbred, and found relatively low numbers of alleles per locus, less heterozygosity than expected, and comparatively high APS, all of which indicate comparatively poor genetic diversity (Cunningham et al., 2001). For the population of Thoroughbreds between 1987 and 1996, an average inbreeding coefficient of 13% was calculated, indicating the presence of significant inbreeding. However, the level of APS in the Thoroughbred was lower than would have been expected under random mating, suggesting that close matings were being avoided. This combination of findings suggest that the Thoroughbred is at risk of losing genetic diversity, and suffering the ensuing adverse effects, if the issue of inbreeding is not addressed and monitored in future.

1.3.2 Genetic diseases of the domestic horse

The domestic horse has 64 (2N) chromosomes, and the genome is diploid, meaning that each cell carries two copies of each gene. Many deleterious traits are encoded on alleles that are said to be ‘recessive’, i.e. their effects are masked by other more ‘dominant’ alleles, such that clinical manifestation of the trait is only observed when the individual possesses two copies of the allele (homozygous recessive). Diseases that are determined by a single gene like this are said to be ‘monogenic’, and can arise in any population. The domestic horse is recognised to suffer from a number of important monogenic diseases. Inbreeding leads to an increase in homozygosity in a population, thus emergence of previously ‘hidden’ recessive traits tends to increase. Many methodologies including candidate gene approaches, association analyses and homozygosity mapping have enabled the identification of the genes responsible for most of the important monogenic diseases of

the horse (Bannasch, 2008). Genetic tests are now available for many of these conditions, so that in theory, no horse should be born with these conditions in future if breeders employ a suitable test on candidate sires and dams pre-mating.

As in human medicine, most common equine genetic diseases are not coded for by a single gene, but are instead the product of a complex interaction between many genes of small effect, and environmental influences (polygenic, complex diseases) (Brosnahan et al., 2010). Since the horse genome was sequenced in 2007, research into these challenging conditions has accelerated. Much has been learned about musculoskeletal conditions such as osteochondrosis (Wittwer et al., 2008; van Grevenhof et al., 2009; Jonsson et al., 2011), lower airway diseases (Ramseyer, 2007; Swinburne et al., 2009; Shakhsi-Niaei et al., 2010), and other traits which are deemed important to breeders such as conformation, coat colour and temperament, but much of this work has neglected the Thoroughbred in favour of other breeds (Oki et al., 2007; Oki et al., 2008; Rieder et al., 2008; Mittmann et al., 2010; Schroderus and Ojala, 2010; Stewart et al., 2010; Tozaki et al., 2012). Investigation of these multifactorial traits is complex, involving the dissection of sources of variation into genetic and environmental parts. Progress in understanding their underlying mechanisms is therefore somewhat slow, but is gathering pace with the advent of powerful, flexible, commercially available software.

1.3.2.1 The genetics of fracture

In human research, studies of the genetics of fracture have focussed on osteoporosis and the inheritance of bone mineral density (BMD), and how these are related to fracture (Duncan et al., 2011). BMD has been shown to be highly heritable in humans (Duncan et al., 2011). The fractures occurring in osteoporosis arise due to weaknesses conferred by abnormally low BMD, rather than the increased BMD observed in equine stress fractures (Shi et al., 2011).

In horses, no previous studies (at the time of writing) have been published on fracture genetics. Alterations in BMD have been shown to occur in Thoroughbreds as a result of training, and this increase has been shown to occur at sites of fatigue fracture (Lanyon, 1987; Riggs, 2002; Entwistle et al., 2008; Shi et al., 2011). The heritability of BMD in the racing Thoroughbred is yet to be reported.

1.3.2.2 The genetics of musculoskeletal conditions

1.3.2.2.1 Osteoarthritis and osteochondrosis

Osteochondrosis (OC) is a disorder of ossification that leads to flattened bone contours, focal necrosis and bone fragments in the joints of many species. When bone fragments are free within the joint space, the condition is called Osteochondritis Dissecans (OCD). It has long been known that OC (and OCD) has a multifactorial aetiology, with diet, growth rate, sex and genetics all playing a part (Philipsson et al., 1993; Ytrehus et al., 2007; Jonsson et al., 2011). In the horse, genetic studies of OC have tended to focus on non-Thoroughbred breeds, and have varied greatly in population sizes and prevalence of OC. Heritability estimates for OC ranged from 0.09 to 0.34 in Standardbreds (Philipsson et al., 1993), 0.05 to 0.13 in Swedish Warmbloods (Jonsson et al., 2011), and was reported as 0.23 in Dutch Warmbloods (van Grevenhof et al., 2009). One study of the Italian Maremmano breed found similar estimates of OC to those listed above, but none were found to be statistically different from zero (Pieramati et al., 2003). A number of studies have found that the heritability of OC varies by joint (van Grevenhof et al., 2009; Jonsson et al., 2011). The magnitude of these heritability estimates suggest that targeted breeding strategies could be of benefit in reducing OC incidence in the populations in question. A number of single nucleotide polymorphisms (SNPs) have been identified that were significantly associated with OC in a Southern German Warmblood population, amongst other breeds, and one study identified a single genome-wide significant SNP on ECA3 in the Thoroughbred (Wittwer et al., 2008; Corbin et al., 2012). There has been limited agreement between studies of OC-associated SNPs in the horse, partly due to a lack of power as a result of small sample sizes. Definitive diagnosis of OC relies on arthroscopy, an invasive and costly procedure, which is not performed in all suspect cases, therefore a degree of misclassification bias could be occurring in studies which rely upon radiographic diagnosis alone, or whose 'control' animals have not also been subjected to arthroscopy to confirm the absence of disease. Larger studies with more precise inclusion criteria are warranted to allow confidence in relevant SNP identification, and studies may need to focus on individual breeds, as these have been shown to differ in their SNP-profiles for OC (Corbin et al., 2012).

Following OC, inflammation is often incited in the affected joint, which can progress to OA. This is only one possible aetiology of OA in the horse, but many different causes are possible. OA has been extensively studied in the human. Numerous associated genes have

been identified, some of which are specific to different ethnic groups, or OA of different anatomical locations (Meyer et al., 1998). Heritability has been estimated from twin studies to be between 0.39 and 0.65. Human OA contains a number of different pathophysiological disorders, based on radiographic, clinical and sometimes surgical definitions. It appears that these different 'subphenotypes' are conferred by different sets of causative genes, which are beginning to be unravelled by genetic studies (Meyer et al., 1998). OA in dogs is a known sequel to joint instability arising due to heritable orthopaedic disorders such as hip and elbow dysplasias, and is often not studied as a standalone trait (Malm et al., 2008; Lewis et al., 2010). The difficulties encountered in attempting heritability studies of OA in dogs include the phenotypic diversity of OA in terms of joint(s) affected, severity of cartilage erosion, radiographic appearance and degree of clinical compromise, as well as the genetic diversity of the species. The prevalence and appearance of OA varies across breeds, meaning that very large sample sizes would be required to estimate OA heritability whilst taking breed and OA phenotype into account. Hip and elbow dysplasia are conditions that are not associated with equidae, thus genetic studies of OA in dogs are not applicable to horses. Genetic studies of OA in horses are lacking.

1.3.2.2.2 Tendon conditions

Oki *et al.* (2008) published one of the only studies of the genetics of equine tendon injury, based on Thoroughbred racehorses in Japan (Oki et al., 2008). This study focused on injuries to the SDFT. Flat racing Thoroughbreds aged between 2 and 5 years were studied, and data were analysed using threshold animal models with a 3-generation pedigree. The heritability of SDFT was estimated to be 0.17-0.19 (s.e. 0.05), which the authors deemed 'moderate' and modifiable through appropriate breeding strategies. No studies have followed to identify the genes responsible for this heritability. Reports of investigations into the heritability of other tendon conditions in other species are sparse, and likely not to be applicable to the horse due to its unique anatomy.

1.3.2.2.3 Ligament conditions

Ligament injury heritability studies in animals tend to be focused on cranial cruciate ligament rupture, and its predisposition within certain dog breeds (Wilke et al., 2006). The heritability of cranial cruciate ligament rupture in Newfoundland dogs was estimated to be

moderate at 0.27. Unlike cruciate ligament rupture in dogs, ligament injuries of the horse tend to be more chronic and degenerative in nature, suggesting a very different underlying aetiology in this species, thus it is likely that the genetics of these conditions are very different. To date, no studies of the genetics of ligament injuries in racing Thoroughbreds have been published. The prevalence of these injuries in the population, coupled with their economic impact on the racing industry, means that genetic studies are warranted to gain a fuller picture of their aetiology in order to reduce their incidence most effectively.

1.3.2.3 The genetics of EIPH

Only one study in recent times has attempted to quantify the effects of genetic influences on the occurrence of EIPH-related epistaxis in horses (Weideman et al., 2004). The presence of blood at the nostrils of 51465 Southern African Thoroughbreds post-exercise was noted and assumed to be due to EIPH, and binary data were analysed using generalized linear mixed models. The heritability of epistaxis as related to EIPH was found to be 0.40 using a logistic sire model, and 0.23 using a logistic animal model, both of which were significantly different from zero (Weideman et al., 2004). The authors concluded that the heritability of epistaxis as related to EIPH in this population was of a magnitude whereby targeted breeding strategies could make fairly swift progress in reducing disease incidence. No other studies of the genetics of EIPH were found in the literature at the time of writing. Future studies are warranted to examine the genetic and phenotypic relationships between epistaxis and EIPH in the Thoroughbred, and to generate estimated breeding values for breeding animals so that genetic gain can be achieved. However, this would require major cooperation from the Thoroughbred industry, as the prevalence of the condition is high.

1.3.3 Genetic concepts and their applications

1.3.3.1 Heritability

Heritability is a key concept in animal breeding. It is defined as the proportion of variance between individuals which is accounted for by genetic differences. More specifically, heritability is additive genetic variance divided by the total phenotypic variance, both of which are subject to change both within and between populations. Total phenotypic

variance is defined as the variation in a trait after accounting for variance attributable to known fixed effects such as age and sex, and it is composed of both genetic and environmental parts;

EQ. 1 $$\sigma_p^2 = \sigma_g^2 + \sigma_e^2$$

where σ denotes variance, and *p, g* and *e* refer to phenotype, genotype and environmental parts, respectively (Falconer and Mackay, 1996). Both genotypic and environmental parts of this equation can be further subdivided. Genotypic variance encompasses the additive effects of genes, as well as dominance and epistasis, i.e. the effects of genes at the same locus, and interactions between genes at different loci. Environmental variance in this sense means all variation not due to genetic influences, including the effects of being from a certain litter, for example, or the plain of nutrition experienced, as well as measurement error and individual stochastic effects. These latter two are known as the residual error. Thus, the equation above can be expanded as follows;

EQ. 2 $$\sigma_p^2 = \sigma_a^2 + \sigma_d^2 + \sigma_i^2 + \sigma_f^2 + \sigma_e^2$$

where subscripts *a, d, i, f* and *e* refer to additive genetic, dominance and epistatic genetic effects, fixed environmental effects and residual error, respectively (Falconer and Mackay, 1996). In reproduction, genes rather than whole genotypes are transmitted from parents to offspring, thus the effects of dominance and epistasis are not inherited. For this reason, 'narrow-sense' rather than 'broad-sense' heritability is used, that is to say that only the additive genetic variance, as a proportion of total phenotypic variance, accounts for differences between individuals, and not genetic variance as a whole. Narrow-sense heritability can therefore be written as (Falconer and Mackay, 1996);

EQ. 3 $$h^2 = \frac{\sigma_a^2}{\sigma_p^2}$$

As can be seen from the definitions above, both the numerator and denominator of the definition of heritability can change, both spatially (between populations) and temporally (within populations). Thus theoretically, heritability is a feature of a specific population at a specific time, and should not be extrapolated to other scenarios. However, in practice, estimates of heritability for traits are remarkably similar across populations and through

time, such that it becomes a very useful parameter for comparison of the relative genetic contribution to variation between different traits or diseases (Visscher et al., 2008).

Low heritability implies that only a small proportion of variation between individuals for the trait in question is determined by additive effects of genes, and the environment plays a greater part. However, this does not indicate that selection for genetic improvement will be fruitless. Response to selection is determined by the additive genetic variance and the correlation between phenotype and genotype (*h*, the square root of heritability). Heritability of a trait may be modest, but the additive genetic variance may be large, thus selection response could be more favourable than the heritability alone would suggest. High heritability means that the phenotype is a good predictor of the genotype. In this scenario, it is important not to assume that phenotype is determined by genotype alone, as environmental factors can always modify the phenotype. Heritability can be manipulated by altering the contributions of the parts of the denominator to the overall phenotypic variance. For example, a reduction in measurement error through more precise recording (e.g. disease diagnosis of orthopaedic problems based on radiography or computed tomography rather that clinical examination) would reduce this source of error and heritability would increase. Similarly, reducing environmental variation more globally, by ensuring as much similarity between individuals in factors such as housing, feeding and management, would allow an increase in heritability through a reduction in phenotypic variance.

The numerator in the definition of narrow-sense heritability is also known as the 'breeding value' (BV). The BV of an animal is the sum of the average effects of the individual's genes which give rise to the mean genotypic value of their offspring. Estimated BVs (EBVs) have been used extensively in livestock industries to rank breeding individuals to ensure maximal response to selection. The EBV of all animals in a pedigree, irrespective of whether they have procreated yet or not, can be obtained using modern statistical techniques.

Historically, heritability was estimated using studies of regression of offspring on parents, twin studies and correlation between half and full-siblings. These balanced designs were superseded by linear mixed model methodologies, capable of incorporating information on every type of relationship within a pedigree, and which are more robust with unbalanced data.

1.3.3.2 Genetic correlations

Pleiotropy is the phenomenon whereby a gene influences the expression of a number of traits rather than just one. Pleiotropy, along with linkage of genes, gives rise to the concept of genetic correlation, where traits are related (either positively or negatively) through their genes. This is distinct from phenotypic correlation, which relates to traits which are significantly more likely (or less likely) to occur and be outwardly detectable in an animal's phenotype. Genetic correlation is important to consider, because it will determine, to some extent, how selection based on one trait might influence the correlated trait. For example, if trait A were positively genetically correlated with trait B, and we selected for trait A in the population, we may affect an increase in trait B concurrently, which could be desirable or deleterious depending on the nature of the traits in question. Phenotypic correlation combines information on both genetic, and environmental correlations. If the heritabilities of the traits are low, then the phenotypic correlation is mostly determined by the environmental correlation, but if heritabilities are high, the genetic correlation is more informative. If trait A was more heritable than trait B, but of less importance in an economic or welfare sense, and these traits were positively genetically correlated, selection based on the phenotype of A (which, due to high heritability, is a reasonable predictor of genotype) could progress and improvements in trait B might be achieved without having to rely on the somewhat uninformative phenotypic expression of trait B. Negative genetic correlations suggest that selection for an increase in one trait will effect a reduction, or *vice versa*, in the correlated trait. Genetic correlations are therefore important considerations in breeding programmes, and must be examined in order to maximise genetic gains, and avoid unintended consequences.

1.3.3.3 Genomic selection

Genomic selection (GS) is a relatively new method devised by Meuwissen et al. (2001) to estimate breeding values (Meuwissen et al., 2001). GS offers potentially large improvements in the accuracy of breeding values over traditional EBVs. The method involves genotyping and phenotyping a large 'training dataset' of animals, so that the effects of each SNP in linkage disequilibrium (LD) with a quantitative trait locus (QTL) are estimated. Following this, young animals from the breeding population need only be genotyped, and the effects of their SNPs summed, to find their genomic estimated breeding value (GEBV). In this way, breeding animals need not be phenotyped, allowing a large

reduction in the generation interval for traits that can only be measured in mature animals. GS has been implemented in dairy cattle, and has been studied as a possible route to faster genetic gain in performance of sport horses (Hayes et al., 2009; Haberland et al., 2012). The accuracies of GEBVs depend upon the linkage between SNPs and QTL, the size of the reference population, the heritability of the trait and the distribution of QTL effects (Hayes et al., 2009). In a lowly heritable trait, many records are required to achieve GEBV of sufficient accuracy; however, Haberland et al. (2012) showed that GS can be of use in selection against a trait with a heritability of 0.15 (Haberland et al., 2012). Another benefit of GS over traditional EBV ranking is that inbreeding appears to increase more slowly under GS than using 'best linear unbiased prediction' (BLUP) estimation, although this benefit needs to be balanced against the large reductions in generation interval that may be possible (Daetwyler et al., 2007). GS could be feasible in the Thoroughbred for conditions of moderate or low heritability, and may offer a route to faster genetic gain versus traditional EBV use, through a reduction in the generation interval.

1.3.4 Aims and objectives

This study was conducted with the following aims:

- Identify important musculoskeletal diseases of the Hong Kong and UK Thoroughbred populations with the available data
- Assess the presence/absence and magnitude of heritability for each condition
- Determine the best type of generalized mixed models to use for each of these analyses
- Examine the possible environmental influences on each condition using the available data
- Examine whether genetic correlations exist between diseases
- Identify any trends over time in genetic disease risk in the populations studied
- Compare and contrast the findings from each population

2. CHAPTER II. MATERIALS AND METHODS

2.1 Datasets

Two datasets were available for use in this study; one was supplied by the British Horseracing Authority (BHA), and one by the Hong Kong Jockey Club (HKJC). They will be referred to as the BHA data and the HK (or HKJC) data, respectively.

2.1.1 Hong Kong data

2.1.1.1 Data source

The data used in this part of the study were provided by the Hong Kong Jockey Club (HKJC). All Thoroughbreds under the care of the HKJC are housed together at Sha Tin Racecourse in the New Territories and managed as a unit. As there are no stud facilities in Hong Kong, all HKJC Thoroughbreds are imported. All racing is on two flat courses: Happy Valley in Hong Kong city, and Sha Tin, in the New Territories. Racing surfaces are composed of all-weather 'dirt' or sand based turf. The HKJC employs a team of full-time veterinarians who are responsible for the clinical care of all horses. Compulsory retirement is enforced for horses that reach 11 years old, have thee officially recorded episodes of EIPH or have three occurrences of cardiac arrhythmia. Retirement for other medical conditions, or for managemental reasons, at the discretion of the veterinarian, trainer or owner, can occur at any time. Horses are continually imported and retired throughout the racing year, with the racing season taking place from September to June. The health and pedigree records used in this study were collected by the HKJC and stored in a purpose-built Microsoft Access database. During the period in which the data for this study were collected there were a median of 1028 horses per annum in training. Disease information was contained in two tables; one contained records of the results of Official Veterinary Examinations (OVEs) carried out between 1995 and 2010 (14886 records), and one table contained free text entries detailing the reason(s) for retirement from racing of all horses retired between 1992 and 2010 (5521 records). OVEs are coded as 'pass', 'fail' or 'request'. A 'request' is a record of the reason that the veterinarian, owner or trainer has examined the horse. Following a 'request', a horse can be passed, failed, or retired. A 'fail' indicates that the horse is not sufficiently healthy to continue racing, and must pass a subsequent OVE before it can be allowed to return to racing (or it may be retired). Many 'pass' records contain little or no extra health information, therefore they were excluded from the OVE data records summary below, but were included in the genetic analyses to

capture the maximum amount of health information. Only retirement and OVE records created from 1st September 1996 onwards were used, due to significant amounts of incomplete information prior to that date. Pedigree information included the name of the sire, dam, paternal and maternal grandsires for each horse. Missing or erroneous parental pedigree information was corrected using the Thoroughbred pedigree information site www.pedigreeinfo.com.

2.1.1.2 Content analysis

Both the OVE and retirement tables were uploaded into content analysis software WordStat v6.1 (Provalis Research). Content analysis procedures were guided by Lam et al. 2007, with some modifications to suit the current data. Following preliminary examination of a proportion of free-text records, a number of categories of disease (or reason for retirement) were defined by the creation of dictionaries listing words or phrases thought to be specific to that disease (or reason). Initial categories included; fracture, tendon/ligament injury, bleeding, poor performance, old age, voluntary retirement, musculoskeletal condition, medical condition, and management decisions. Further dictionaries were subsequently created to further subdivide these categories, e.g. sesamoid fracture, ligament injury, EIPH, osteoarthritis. Each word or phrase was verified as pertaining to its disease category by assessment of the 'keyword in context' strings. Uncategorized records were examined, and updates were made to dictionaries to include spelling errors and synonyms in an iterative process until all records had been categorized. Dictionaries used to categorize OVE and retirement records differed slightly due to different terminologies used in each. What follows is a brief summary of the OVE and retirement data considered separately.

2.1.1.3 OVE records summary

After exclusion of 'pass' OVEs, 8762 records remained, which were from 3516 horses, 99.4% of which were male. The number of OVEs per horse ranged from 1 to 16 (mean 2.49, median 2, mode 1). Forty six trainers were assigned to the 3516 horses, with each horse having OVEs under up to 4 different trainers (mean 1.23, median 1, mode 1 trainers per horse). Each trainer trained between 3 and 245 horses (mean 92.28, median 98, mode 113 horses per trainer) over the period covered by the data.

Where OVE or retirement categories could overlap, for example with ‘musculoskeletal’ and ‘fracture’, the categories are mutually exclusive, so that fracture would be excluded from musculoskeletal diagnoses and *vice versa*. The number of OVEs reported annually was proportional to the number of horses at risk (housed and in training with the HKJC) (Figure 2-1). The greatest proportion of OVEs were of four year-old horses, and most horses retired at age 5 years (Figures 2-2 and 2-6). The most common cause of OVEs was a musculoskeletal problem (not including lameness, DJD etc.), followed by lameness (Figure 2-3).

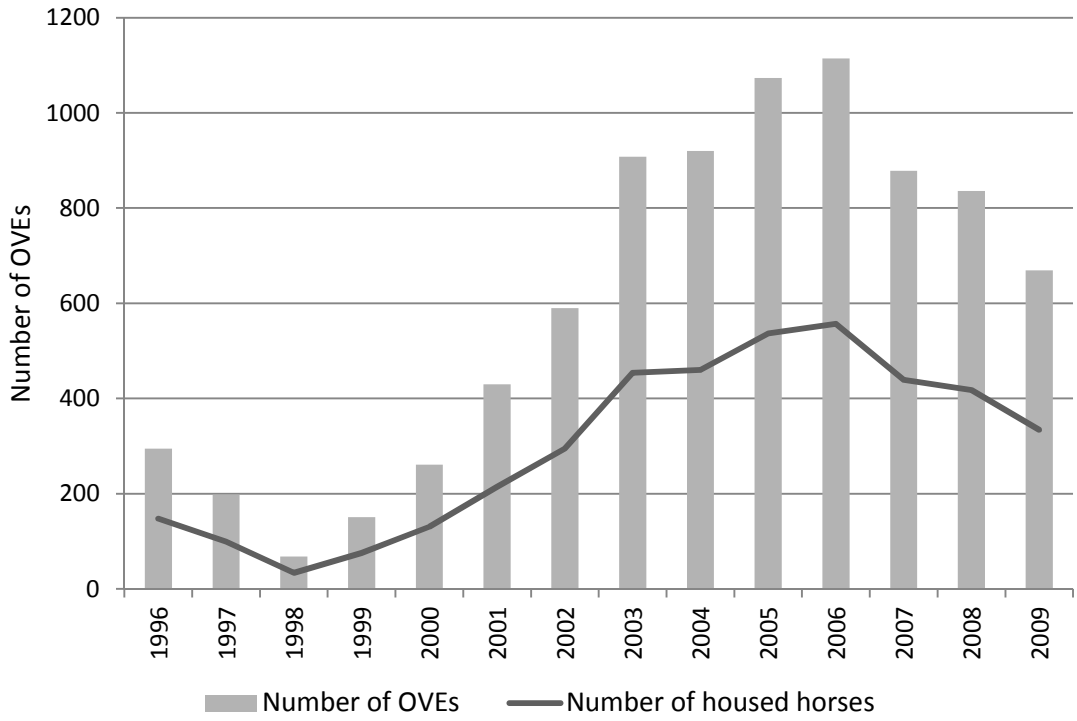

Figure 2-1. The number of OVEs reported yearly from 1996 to 2009, and the mean number of horses under the care of the HKJC over that time.

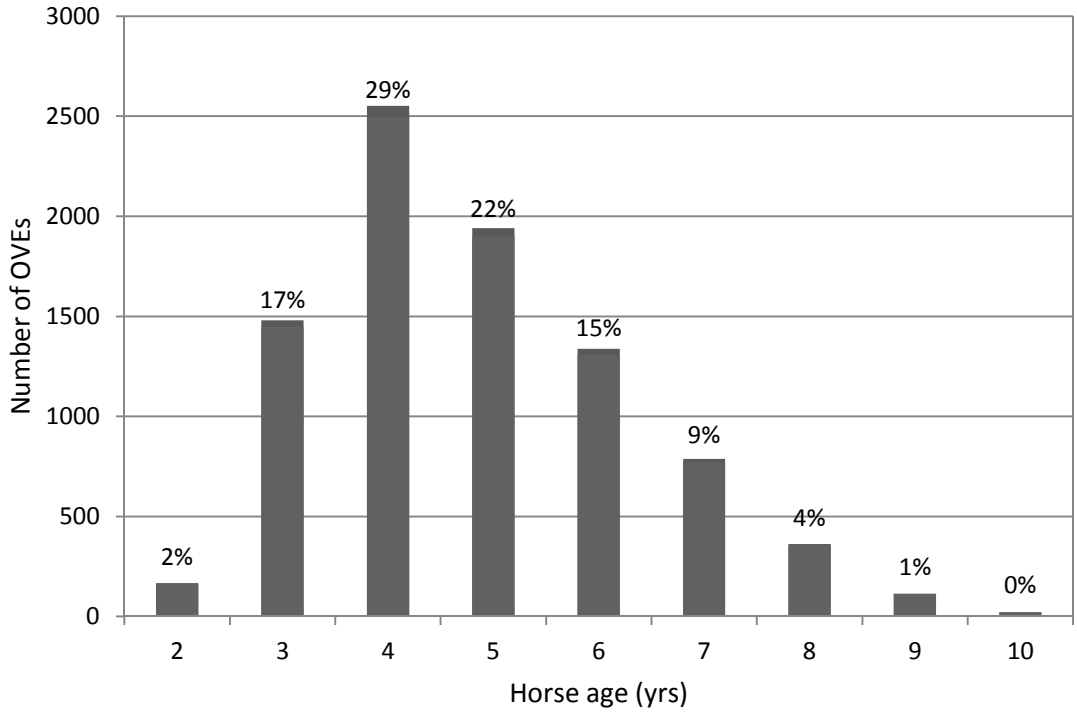

Figure 2-2. The number (and percentage) of OVEs for horses of between 2 and 10 years old. Percentages are the number of OVEs of that age group divided by the total number of OVEs (n=8762).

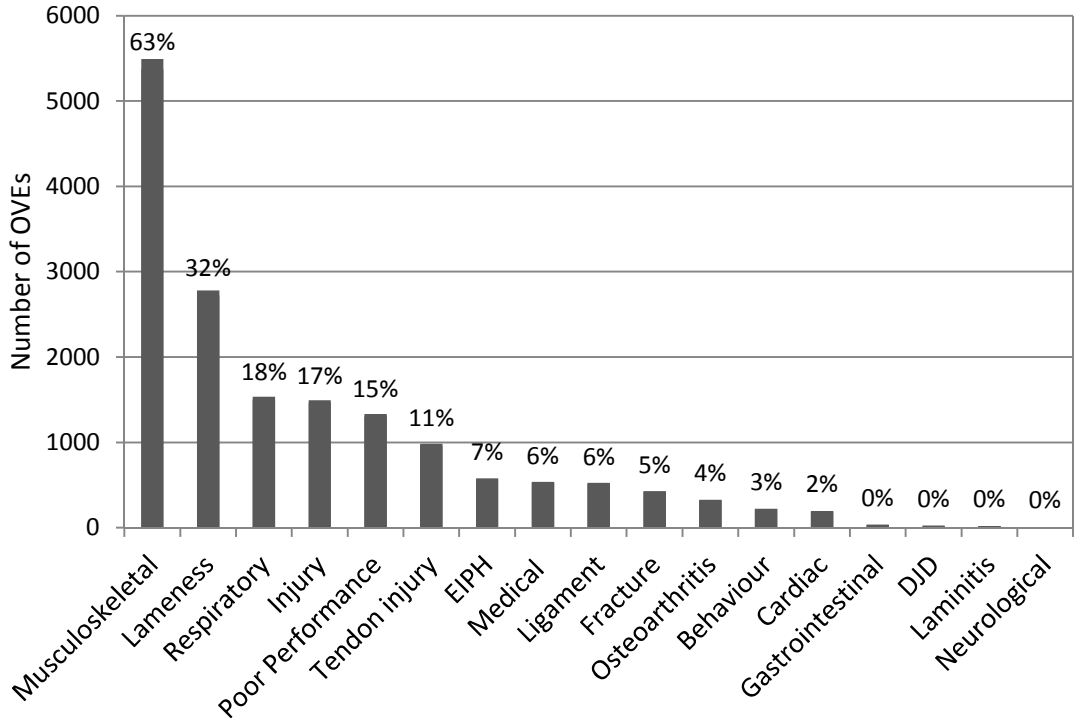

Figure 2-3. The number (and percentage) of OVEs that were reported to be due to the listed categories. An OVE may report more than one cause, therefore percentages exceed 100%. All categories are mutually exclusive.

Figure 2-4 gives the number of fractures of various anatomical sites diagnosed in Hong Kong over the study period.

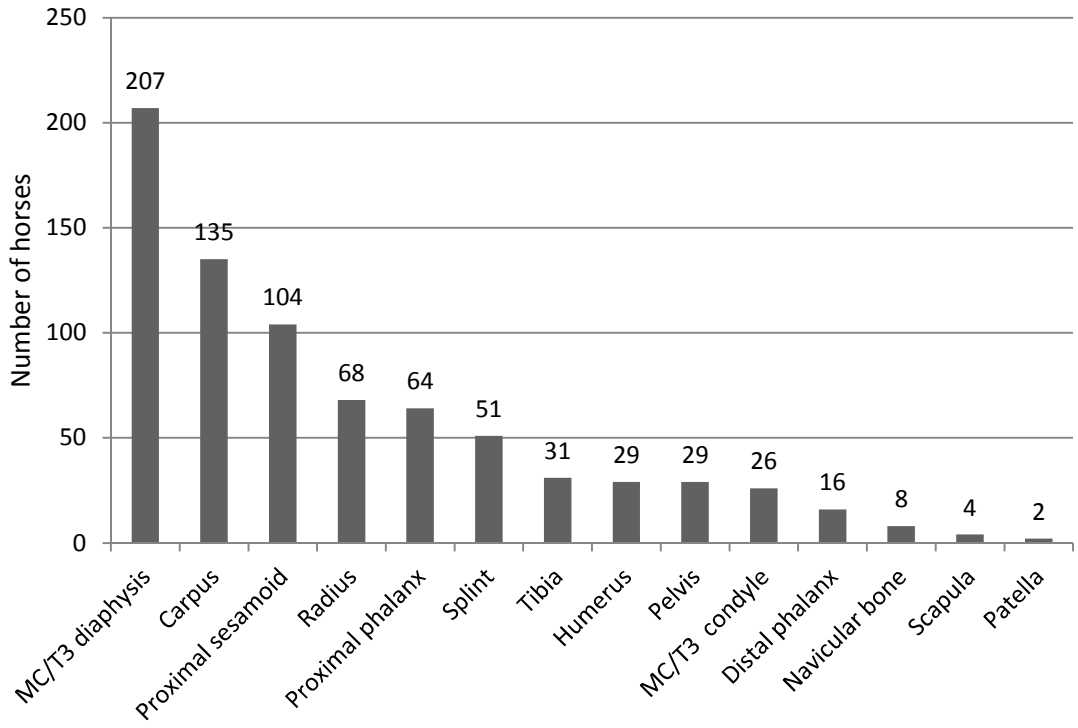

Figure 2-4. The number of horses which had at least one OVE reported for the listed fracture types. Multiple OVEs for different fractures were possible, as were multiple fractures diagnosed at a single OVE. MC/T3 stands for the third metacarpal or metatarsal bone.

2.1.1.4 Retirement records summary

There were 5089 retirements from racing in the period between September 1996 and December 2010. Almost all (99.3%) of these horses were male. The annual number of retirements has stayed stable at around 350 horses over the study period (Figure 2-5). For the complete years between 1997 and 2010 the annual number of retirements ranged from 315 to 450. Horses were retired from racing between the ages of two and eleven, and the most common age at which horses retired was five years old (Figure 2-6). Most retirements were on medical grounds (veterinary diagnoses excluding those included in other categories in Figure 2-7, e.g. cardiology, dermatology), with musculoskeletal conditions being the second most prevalent retirement reason (Figure 2-7). Seven percent of OVEs were for EIPH/epistaxis, and eight percent of retirements stated bleeding as a cause (due to terminology difficulties, EIPH could not be specifically distinguished from other bleeding problems in the retirements data, as some retirement records were simply labeled as ‘bleed’, ‘bled’ or ‘bleeder’) (Figures 2-3 and 2-7). The most common fracture reported at an OVE was fracture of the MC/T3 diaphysis, however proximal sesamoid bone fractures were more commonly a cause of retirement from racing (Figures 2-4 and 2-8).

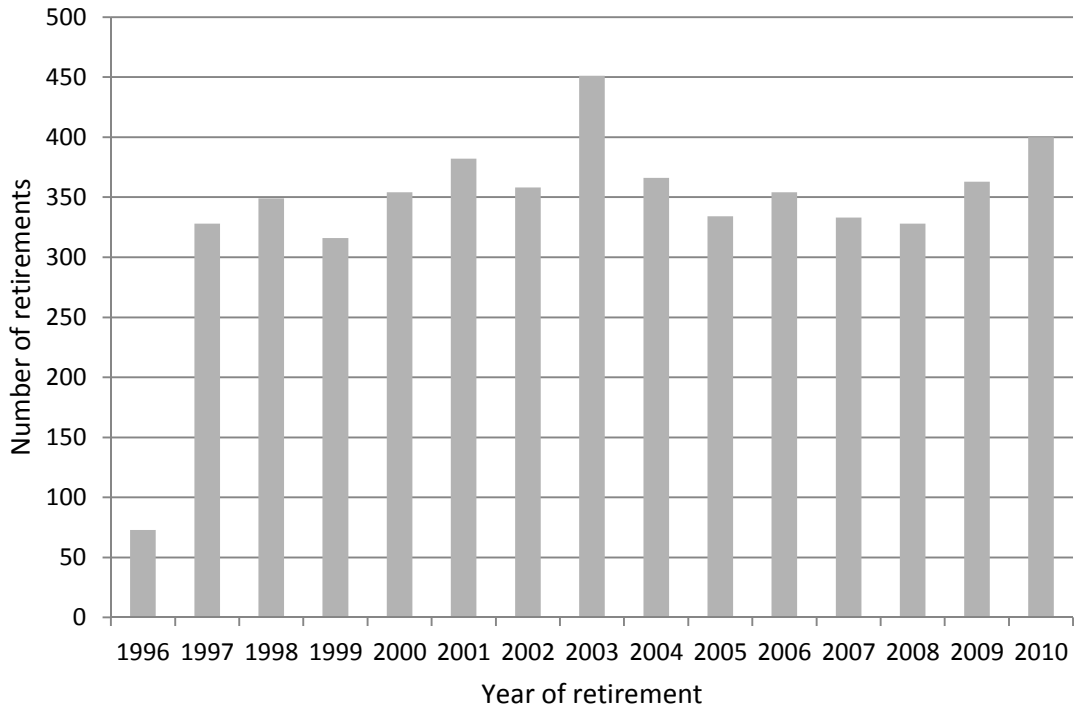

Figure 2-5. The number of horses retiring annually from the HKJC, between 1996 and 2010. The first year shown (1996) was restricted to retirements after 1st September.

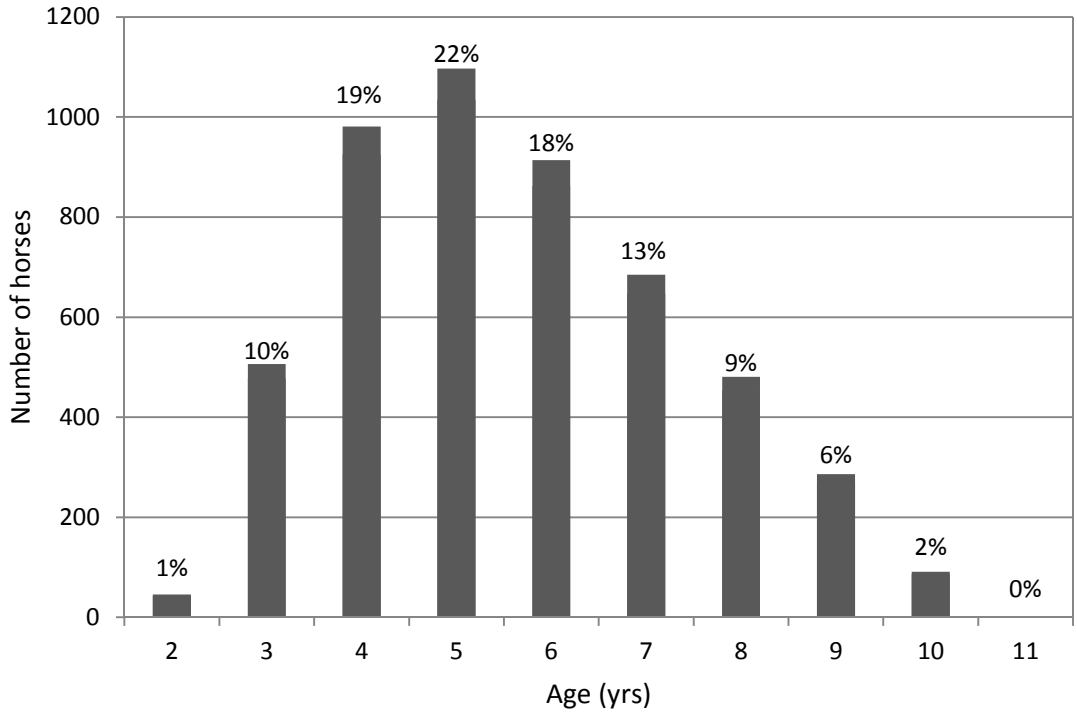

Figure 2-6. The number (and percentage) of horses that were retired from the HKJC between Setpember 1996 and December 2010, that were of different ages (in years).

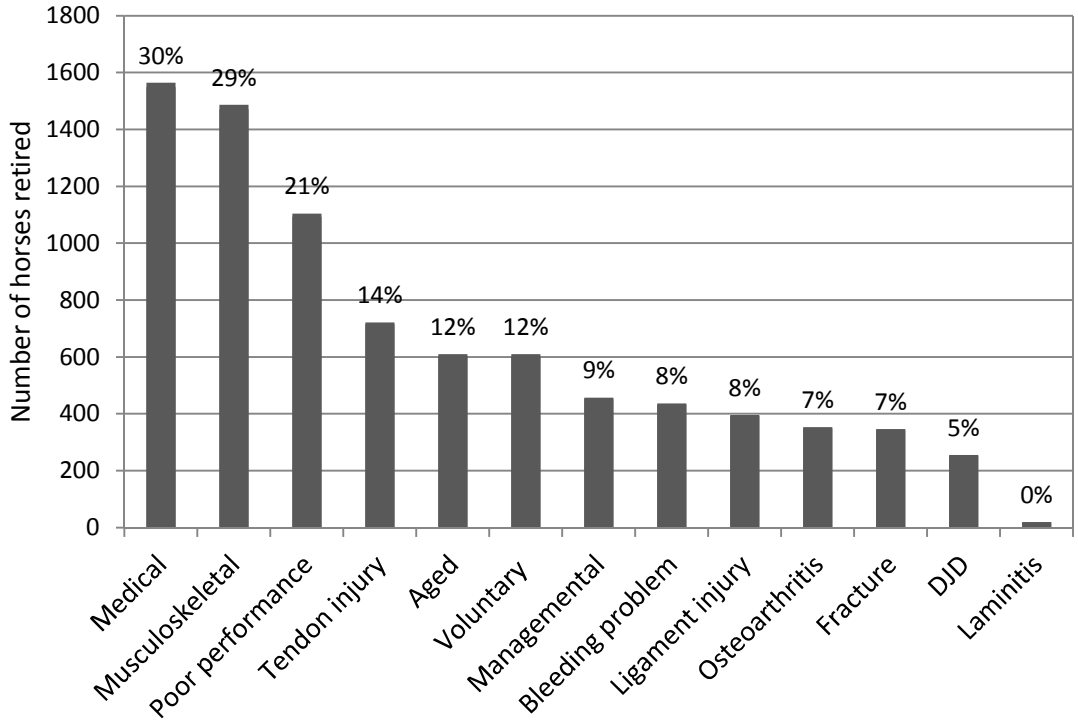

Figure 2-7. Number (and percentage) of retirement records that reported the listed reasons for retirement. Multiple reasons for retirement were possible, thus percentages do not sum to 100%. All categories are mutually exclusive.

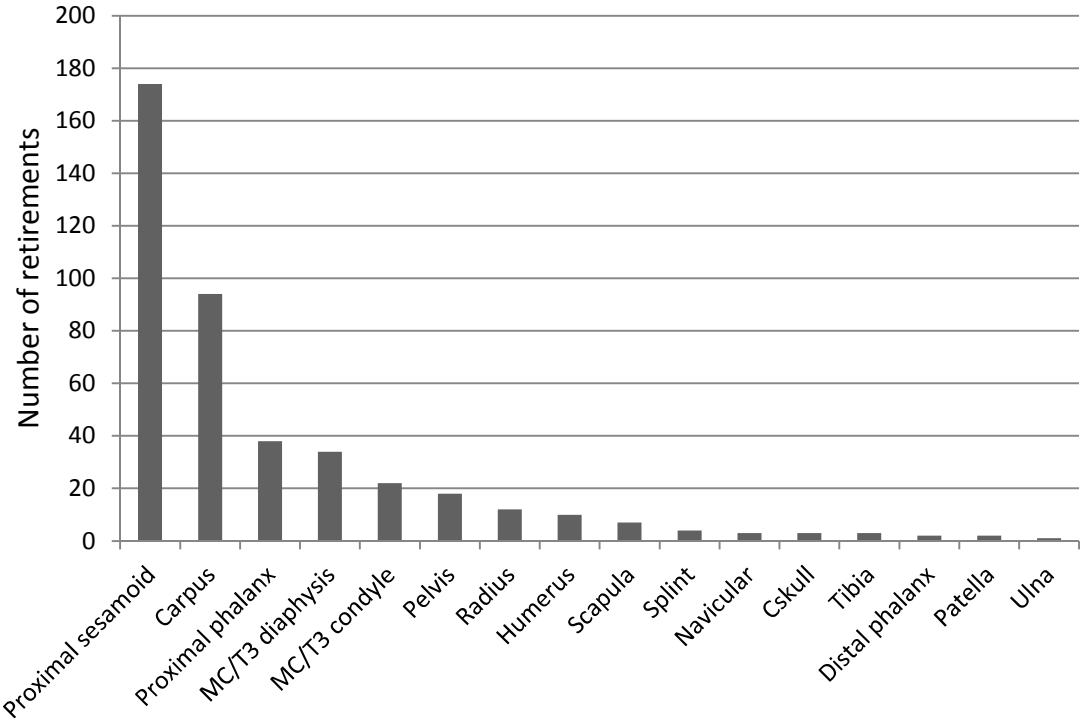

Figure 2-8. The number of retirements that were reported to be due to the listed fracture types. Multiple fractures in a single retirement record were possible. MC/T3 stands for the third metacarpal or metatarsal bone.

2.1.1.5 Variables and summary of merged HK datasets

The OVE and retirement data were merged after indexing by alphanumeric horse identification codes, and only horses that had retired were included in this final dataset. The final dataset therefore comprised all horses that retired from racing over the period of the study and their full OVE history during their time in Hong Kong. The presence or absence of each disease category (coded 1 or 0) in each horse was exported as a Microsoft Excel spreadsheet. In this final dataset, each horse was included only once, thus its 'disease status' was 1 for a condition if it had ever been diagnosed with that condition, whether during an OVE, or as a reason for retirement, or both. Repeated diagnoses of the same condition were not used. All text or alphanumeric variables were numerically recoded before heritability analyses were performed. The final HKJC data contained records of 5062 horses.

After cleaning, the final HKJC data contained 12 variables. From these, a further 10 variables were created in Microsoft Excel, as described in Table 2-1. Table 2-2 describes the continuous variables available within the HKJC dataset, and Table 2-3 lists the number of horses that were imported from different countries. The horses in the final dataset were

born between September 1985 and November 2007, were imported into Hong Kong between July 1989 and July 2010, retired between September 1996 and December 2010 and included 99.3% males and 93.6% neutered animals. These horses descended from 1294 sires and 4553 dams. Each sire contributed between 1 and 167 offspring to the final dataset (mean 3.9, median 2, mode 1) and each dam contributed between 1 and 5 offspring (mean 1.1, median 1, mode 1).

Table 2-1. HKJC data variables with their derivations, where appropriate. Blank cells indicate that a variable was available and used unchanged from the original data.

Variable Name	Source	Description
Date of retirement		Microsoft Excel date format, e.g. 40909 (01/01/2012)
Year of retirement	Date of retirement	Four-digit number e.g. 1999
Age at retirement		Years, range 2 to 11
Sex	Traditional terms e.g. colt, gelding, converted to male or female	Male 1, female 0
Neuter status	Extracted from traditional sex definitions	Neutered 1, not neutered 0
Country of origin		Coded numerically, alphabetically, from 1 to 16
Continent of origin	Country of origin	Coded numerically from 1 to 6
Hemisphere of origin	Country of origin	Northern 1, Southern 2
Sire country of origin		Coded numerically, alphabetically, from 1 to 16
Dam country of origin		Coded numerically, alphabetically, from 1 to 16
Sire of dam country of origin		Coded numerically, alphabetically, from 1 to 16
Date of birth		Microsoft Excel date format, e.g. 40909 (01/01/2012)
Year of birth	Date of birth	Four-digit number e.g. 1999
Date of import		Microsoft Excel date format, e.g. 40909 (01/01/2012)
Year of import	Date of import	Four-digit number e.g. 1999
Stakes won in career		Continuous, HKD*
Number of career starts		Whole numbers, 0 to 112
Trainer at time of retirement		Trainer initials numerically recoded alphabetically, 1 to 46
Date of first race		Microsoft Excel date format, e.g. 40909 (01/01/2012)
Career length in days	Date of retirement minus date of first race	Continuous, 0 to 2826
Career length in years	Career length in days divided by 365	Continuous, ≥ 0
Racing intensity (average number of weeks between races over whole career)	Career length in weeks (career length in days / 7) divided by the number of career starts	Continuous, 0 to 78.14

*Hong Kong Dollars

Table 2-2. Range, mean and median of continuous variables in the HKJC dataset.

Variable	Range	Mean	Median
Age at retirement (yrs)	2-11	5.6	5
Stakes won in career (HKD)	$0\text{-}7.5\times10^{7}$	1.5×10^{6}	8.6×10^{5}
Number of career starts	0-112	22.5	18
Career length (yrs)	0-7.7	2.4	2.2
Intensity (average number of weeks between adjacent races over whole career)	0-78.1	6.15	5.46
Number of OVEs	0-27	2.9	2
Number of horses per trainer at time of retirement*	1-266	110	113

*No trainer is permitted to train more than 65 horses at any one time. These data refer to the total number of horses trained by each trainer over the whole period of the study.

To ensure adequate representation at all levels of each categorical variable, all country of origin variables were grouped as follows: group 1 – Australia and New Zealand, group 2 – UK, Ireland and the USA, group 3 – all others. Countries were also investigated as continent groups: group1 – Europe, group 2 – Australasia, group 3 – North America, group 4 – all others. All horses that retired aged 10 or greater were grouped together, as were horses born in 1985 to 1987, and 2006 to 2007.

Table 2-3. Number of horses imported from different countries between September 1996 and 2010. Parental and grandparental country origins refer to the number of horses whose relative originated from that country, e.g. there were 15 horses whose sires originated from Argentina, but the horse itself may have originated elsewhere.

	Country of origin	Sire's country of origin	Dam's country of origin	Sire of dam's country of origin
	Number of horses			
Argentina	40	15	39	26
Australia	1729	891	1303	632
Brazil	3	1	3	0
Canada	29	86	41	169
China	0	1	2	1
France	75	101	140	292
Great Britain	390	449	436	648
Germany	15	9	10	5
Ireland	715	696	685	735
Italy	1	1	3	2
Japan	5	11	5	4
New Zealand	0	0	0	1
South Africa	42	8	33	15
USA	556	2438	956	2176
Zimbabwe	2	0	0	0

Table 2-4 lists eight important veterinary conditions detected in the HKJC data by content analysis, with the number and proportion of horses that were diagnosed with each condition.

Table 2-4. Number and percentage of horses diagnosed with each of the listed conditions (OVE and retirement datasets merged). Distal limb fractures were a subset of all fractures.

Condition	Number (%) of diagnosed horses (n=5062)
Degenerative joint disease	269 (5.3)
Distal limb fracture	152 (3.0)
EIPH	987 (19.5)
Fracture	677 (13.4)
Ligament injury	585 (11.6)
Osteoarthritis	510 (10.1)
Suspensory ligament injury	524 (10.4)
Tendon injury	957 (18.9)

Pedigree information for the HKJC dataset contained identities of 12169 horses, and up to eight generations of information per horse.

2.1.2 BHA data

Information on all equine injuries reported at all of the UK's sixty racecourses between January 2000 and January 2010 were available from the BHA in a Microsoft Excel spreadsheet (900775 starts). Every race in the UK is compulsorily attended by two racecourse veterinarians (three for jump races), and every equine medical event must be recorded using a Racecourse Veterinary Consultation Form. A BHA Veterinary Official must also be present at every race meeting, and it is the Official's responsibility to upload the medical information into a central BHA database following every meeting. Each medical event is categorized using one or more tick-boxes available from a large set of possible diagnoses in the BHA database. Diagnoses are made using clinical examinations on site, without extensive use of diagnostic aids such as radiography or ultrasonography in the majority of cases. For use in this study, three prevalent conditions were selected for investigation, namely distal limb fracture (inclusive of carpus/tarsus), epistaxis and SDFT injury.

Variables available for use in the BHA data were: date and type of race (flat, hurdle, steeplechase or National Hunt flat), horse name and identity of the parents and grandparents, date of birth, winnings over the lifetime of the horse (British pounds) and sex

(colt, filly, gelding, mare or rig). Sex was converted to male (0) or female (1) for analyses. The number of starts by each horse was derived from the number of records each horse contributed to the dataset. The length of 'career' was calculated by subtracting the date of the first race from the date of the last race plus one (this variable therefore excludes the initial training period before the first race is run). Career length in years was calculated by dividing this value by 365. The 'intensity' of racing was calculated by dividing the number of starts by the career length in days. Career length, number of starts, and intensity of racing must be interpreted with caution as some horses may not have completed their racing careers during the study period. Continuous variables are described in Table 2-5. Race type was used to create a racing 'profile' for each horse's career, which contained either all flat races, all jump races or a mixture of flat and jump races. Racing profile was also investigated by categorizing all horses as either having run only flat races, or not.

During the time in which these data were collected, 84380 races were run over 3464 race meetings, and 78151 horses were represented in the dataset, 63.7% of which were males (49804 of 78151). Of these races, 52254 (61.9%) were flat, 16636 (19.7%) were hurdle, 2161 (2.6%) were NHF, and 13329 (15.8%) were steeplechase races. The size of the field ranged from 2 to 40 horses per race. The mean number of runners per race for flat, hurdle, NHF and steeplechase races were 11.0, 11.2, 12.9 and 8.5, respectively. All horses were born between April 1983 and June 2007, and descended from 3089 sires and 39033 dams. Each sire contributed between 1 and 618 offspring to the dataset (mean 25.3, median 4, mode 1), and each dam contributed between 1 and 13 offspring (mean 2, median 1, mode 1). Each horse contributed between 1 and 177 starts to the data (mean 11.5, median 7, mode 1).

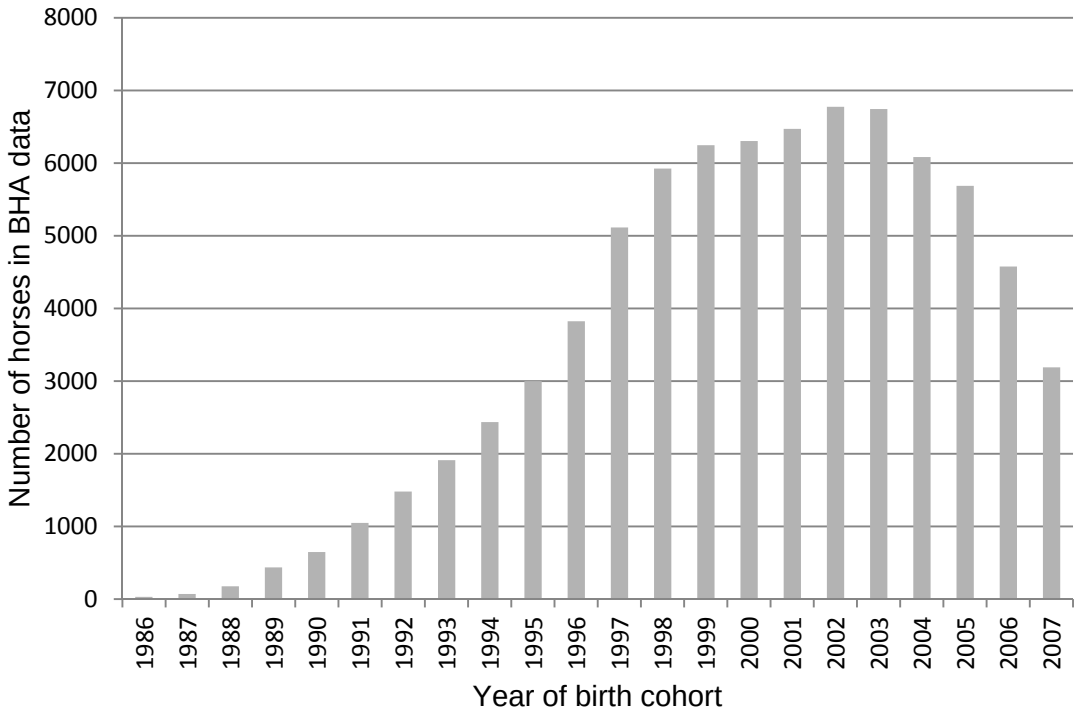

Figure 2-9. The number of horses per year of birth cohort that contributed phenotypic information to the BHA dataset (n = 78151).

The number of horses that only competed in flat races or only competed in jump races were 38016 (48.6%) and 28491 (36.5%), respectively. There were 11644 (14.9%) horses that competed in a mixture of flat and jump races.

Table 2-5. Description of continuous variables in the BHA dataset.

Variable	Range	Mean	Median
Career length (yrs)*	0.00-9.72	1.53	1.02
Number of starts†	1-177	11.53	7
Intensity of racing‡	0.00-1.00	0.14	0.03
Winnings (British Pounds)	$0\text{-}4.4\text{x}10^6$	13702	1841

*Career is defined as the number of years between the first and last race dates. †This was calculated from the available data as the number of records per horse. ‡Intensity is the number of starts divided by the career length in days (last minus first race date, plus 1). Some horses may not have completed their racing careers within the study period, therefore these variables should be interpreted with caution.

The number of horses diagnosed at least once with distal limb fracture, epistaxis or SDFT injury were 797 (1.0%), 1667 (2.1%) and 2070 (2.6%), respectively. The number, and proportion, horses diagnosed with each condition is shown in Table 2-6. Table 2-7 shows the distribution of distal limb fracture diagnoses across all race types.

Table 2-6. The number of horses with one or more diagnoses of distal limb fracture, epistaxis or SDFT injury during, diagnosed at each race type.

	All race types (n=78151)	Flat races (n=49660)	Hurdle races (n=31644)	NHF races (n=13949)	Steeplechase races (n=16390)
Distal limb fracture	797 (1.0%)*	290 (0.6%)	267 (0.8%)	42 (0.3%)	198 (1.2%)
Epistaxis	1667 (2.1%)	592 (1.2%)	599 (1.9%)	23 (0.2%)	520 (3.2%)
SDFT injury	2070 (2.6%)	223 (0.4%)	1110 (3.5%)	55 (0.4%)	709 (4.3%)

*Percentages are calculated as the number of horses with the appropriate diagnosis made during the appropriate race type, divided by the total number of horses that competed in that race type. A single horse may have been diagnosed with a condition in races of more than one type, thus the number of horses diagnosed in different race types may exceed the total number of horses diagnosed with that condition.

Table 2-7. Distal limb fracture rate per race type in BHA data.

	Number of distal limb fractures	Number of starts	Distal limb fracture rate per 1000 starts
All race types	797	900757	0.89
Flat	291	573264	0.51
Hurdle	267	186106	1.43
NHF	42	27920	1.50
Steeplechase	198	113476	1.74

For heritability analyses, the BHA data were used in two ways. All records and variables described above were used to generate repeatability models. This 'full' dataset will be referred to as the 'repeatability' (REP) dataset. These data were also collapsed to produce one record per horse, with disease variables now indicating whether each horse had had one or more diagnoses of each condition over the period of its racing history covered by the data, i.e. binary data (0 indicated a horse had never been diagnosed with that condition, 1 indicated a horse had been diagnosed at least once with that condition). This dataset will be referred to as the 'per horse' (PH) dataset. Variables used in the PH dataset were as follows: sex, date of birth, number of career starts, racing profile, career length, intensity of racing and winnings.

Pedigree information for the BHA dataset included 118138 horse identities, and up to 4 generations of information per horse.

2.2 Model building

2.2.1 Univariable analyses

In both the HK and BHA datasets, the relationship between every combination of explanatory variables and all possible diagnoses were investigated by univariable logistic regression analyses, using R v.2.15.1 (R Development Core Team, 2008). Each explanatory variable was plotted against the log odds of the outcome to assess linearity, and where non-linear relationships were found, suitable transformations were investigated. Variables were retained for multivariable analyses if their *p*-value was <0.2. All retained variables, for each diagnosis, were ordered by log likelihood (highest to lowest) of the univariable model (Dohoo et al., 2003).

2.2.2 Multivariable analyses

Sire or animal random effects variables were used in each model to incorporate genetic relationships. All variables found to be associated with each condition were sequentially added to the appropriate sire or animal linear or logistic regression model in order of descending log likelihood, and retained if their *p*-value was <0.05, and the likelihood ratio test (LRT) *p*-value was also <0.05. The general form of the linear model was:

EQ. 4 $$Y = Xb + Za + e$$

where Y is the vector of observations, X and Z are known incidence matrices, b is the vector of fixed effects, a is the vector of random additive genetic effects with the distribution assumed to be multivariate normal with parameters $(0, \sigma_s^2 \mathrm{I})$ for sire models and $(0, \sigma_a^2 \mathrm{A})$ for animal models, e is the vector of residuals with multivariate normal distribution and parameters $(0, \sigma_e^2 \mathrm{I})$, and where I denotes an identity matrix, A is the numerator relationship matrix, and σ^2 denotes variance (van der Werf, 2012). The general logistic model form was:

EQ. 5 $$\log\left(\frac{p}{1-p}\right) = Xb + Za + e$$

where *p* denotes the probability of the condition in the population, and all other components are as before. Potential correlations between explanatory variables were investigated by generating correlation coefficients between every pair of retained variables in every model. All correlation coefficients of >0.8 were investigated by removal of each of the correlated variables in turn, to assess the impact on the log odds of the remaining variable. All pair-wise interactions between all retained explanatory variables were investigated, and were retained if all variable *p*-values remained <0.05 with a significant likelihood ratio test (*p*-value <0.05) (Dohoo et al., 2003). Finally, the significance of the genetic random variable (sire or animal) to the overall fit of each model was assessed by likelihood ratio tests. All model building was performed in R v.2.15.1 (R Development Core Team, 2008).

2.2.3 Heritability, genetic correlation and EBV estimation

Where inclusion of a sire or animal random effect did not significantly improve the fit of any of the final models (animal or sire, linear or logistic), it was assumed that there was insufficient evidence in the current data to detect heritability, and it was not investigated further. These conditions were degenerative joint disease (DJD) and distal limb fracture in the HKJC data. The heritability of each remaining condition was estimated using Residual Maximum Likelihood (REML) with ASReml v.3 genetic analysis software (VSN International). Residual variance in logistic models was set to $\frac{\pi^2}{3}$. Heritability in sire models was determined by:

EQ. 6 $$h^2 = \frac{\sigma_s^2 \times 4}{\sigma_p^2}$$

where σ_s^2 is sire variance, and σ_p^2 is total phenotypic variance, composed of sire and residual variance (Williams et al., 2001). Animal model heritability was determined by:

EQ. 7 $$h^2 = \frac{\sigma_a^2}{\sigma_p^2}$$

where σ_a^2 is the additive genetic variance.

Genetic correlations were determined using REML by generation of appropriate bivariate sire and animal linear models without fixed effects. The genetic correlation between flat and jump racing populations in BHA data were determined by using bivariate models with flat or jump racing as dependent variables.

EBVs were produced in ASReml during linear animal model heritability estimation. Accuracy of EBVs was calculated as follows:

EQ. 8 $$accuracy = \sqrt{\left(1 - \frac{s_i^2}{(1+f_i)\sigma_a^2}\right)}$$

where s_i is the standard error of the breeding value estimate for individual *i* provided by ASReml, f_i is the inbreeding coefficient of individual *i*, and σ_a^2 is the genetic variance (van der Werf, 2012).

2.2.4 BHA Repeatability

The BHA REP dataset was used to calculate heritability and repeatability of each of the three traits studied. Both animal and sire models were used to calculate repeatability. The general repeatability model form was as follows:

EQ. 9 $$Y = Xb + Za + Zp + e$$

where *p* is the vector of permanent environment effects, assumed to be multivariate normal with parameters (0, $\sigma_{pe}^2 I_p$), I_p is an identity matrix of appropriate order, σ_{pe}^2 is permanent environmental variance, and all other components are as described previously (Williams et al., 2001). In sire repeatability models, heritability was determined by:

EQ. 10 $$h^2 = \frac{\sigma_s^2 \times 4}{\sigma_s^2 + \sigma_e^2}$$

and repeatability was:

EQ. 11 $$r = \frac{4\sigma_s^2 + \sigma_{pe}^2}{\sigma_s^2 + \sigma_{pe}^2 + \sigma_e^2}$$

In animal repeatability models, heritability was determined by:

EQ. 12 $$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_{pe}^2 + \sigma_e^2}$$

and repeatability was:

EQ. 13 $$r = \frac{\sigma_a^2 + \sigma_{pe}^2}{\sigma_a^2 + \sigma_{pe}^2 + \sigma_e^2}$$

3. CHAPTER III. FRACTURE ANALYSES

3.1 Hong Kong fracture analysis results

The HKJC data contained records of 677 (13.4% of all horses) horses that had been diagnosed with a fracture, and of these, 152 (3.0% of all horses) that had been diagnosed with a distal limb fracture.

3.1.1 Multivariable analysis

Table 3-1 shows results of the significant fixed effects in final models of fracture. Only year of birth was significant in the multivariable model of fracture. Increasing year of birth was associated with increased odds of fracture.

Table 3-1. Results of multivariable sire regression models on linear and logistic scales.

	Linear model	**Logistic model**		
	Wald *p*-value	**Wald *p*-value**	**Odds ratio**	**95% CI***
Year of birth	<0.001	<0.001	1.115	1.009-1.114

*Confidence interval (1.96 x standard error).

Inclusion of animal or sire in the final models for distal limb fracture did not significantly improve the fit of any of these models, thus heritability analyses were not conducted on this outcome (Table 3-2). Both sire and animal linear and logistic regression models of fracture were significantly improved by addition of the genetic variable (*p*-value <0.05).

Table 3-2. Likelihood ratio test statistics and *p*-values of final models of distal limb fracture, and fracture. LRT statistic is determined by $2(logL_{full} - logL_{red})$ where *L* is likelihood, and *full* and *red* indicate the final model, and the 'reduced' final model, i.e. minus the genetic random variable only (sire or animal).

Condition		Model Form	LRT statistic	*p*-value
Distal limb fracture	Animal	Linear	0.50	0.48
		Logistic	-242.24	NA
	Sire	Linear	1.78	0.18
		Logistic	-453	NA
Fracture	Animal	Linear	5.28	0.02
		Logistic	382.84	<0.001
	Sire	Linear	7.24	0.01
		Logistic	163.36	<0.001

3.1.2 Heritability and genetic correlation estimation

Table 3-3 shows heritability estimates of fracture with their standard errors. All estimates of fracture heritability exceeded their respective standard error, but all 95% confidence intervals spanned zero.

Table 3-3. Heritability estimates of fracture from animal and sire linear and logistic regression models.

Model form		h^2	Standard error	95% CI*
Sire	Linear	0.0400	0.0248	-0.0086 – 0.0886
	Logistic	0.1087	0.0632	-0.0152 – 0.2326
Animal	Linear	0.0283	0.0187	-0.0084 – 0.0650
	Logistic	0.0470	0.0315	-0.0147 – 0.1087

*Confidence interval (1.96 x standard error).

Table 3-4 contains genetic and phenotypic correlations between fracture and all other 'heritable' conditions analysed in the HKJC dataset. Upper 95% confidence interval limits that exceed 1 are theoretically impossible, but they are included for completeness. Positive significant genetic correlations were found between fracture and ligament injury, OA and suspensory ligament injury. A significant positive phenotypic correlation was found

between fracture and osteoarthritis, and negative phenotypic correlations were found between fracture and EIPH/epistaxis and tendon injury.

Table 3-4. Genetic and phenotypic correlations (with standard errors and 95% confidence intervals) between fracture and the named conditions. Sire model results appear above animal model results, which are in bold type. P-values are based on LRT between the unconstrained model and the same model with the genetic correlation constrained to 0.00001. 'EIPH' includes diagnoses of epistaxis.

	r_g*	s.e.†	95% CI‡	*p*-value	r_p§	s.e.	95% CI
EIPH	0.1784 **0.1430**	0.1963 **0.1906**	-0.2060 - 0.5631 **-0.2306 – 0.5166**	0.35 **0.44**	-0.0536 **-0.0537**	0.0141 **0.0142**	-0.0810 - -0.0260 **-0.0815 - -0.0259**
Ligament injury	0.5660 **0.4193**	0.2671 **0.2655**	0.0425 - 1.0895 **-0.1011 – 0.9397**	0.03 **0.11**	-0.0029 **-0.0027**	0.0141 **0.0141**	-0.0310 - 0.0247 **-0.0303 – 0.0249**
Osteoarthritis	0.8499 **0.8930**	0.1347 **0.1422**	0.5859 - 1.1139 **0.6143 – 1.1717**	<0.001 **<0.001**	0.1682 **0.1662**	0.0138 **0.0138**	0.1412 - 0.1952 **0.1392 – 0.1932**
Suspensory ligament injury	0.5902 **0.4225**	0.2941 **0.2912**	0.0138 - 1.1666 **-0.1483 – 0.9933**	0.05 **0.14**	-0.0051 **-0.0049**	0.0141 **0.0141**	-0.0220 - 0.0225 **-0.0325 – 0.0227**
Tendon injury	0.0570 **-0.1479**	0.2345 **0.2127**	-0.4030 - 0.5166 **-0.5648 – 0.2690**	0.81 **0.50**	-0.1069 **-0.1081**	0.0140 **0.0140**	-0.1340 - -0.0790 **-0.1355 - -0.0807**

*Genetic correlation, †standard error, ‡confidence interval, §phenotypic correlation.

3.1.3 Estimated breeding values for fracture

EBVs were extracted from animal linear and logistic models. Only EBVs for animals that had a recorded date of birth were used. Accuracies ranged from 0.14 to 0.73 from the linear animal model, with a mean of 0.24. Accuracies from the logistic animal model ranged from 0.07 to 0.73, with a mean of 0.20. Figures 3-1 and 3-3 chart the mean EBV and accuracy per year of birth cohort between 1986 and 2007, from linear and logistic animal models, respectively. There was an increasing trend in both mean EBV and EBV accuracy from both models over the specified time period. Figures 3-2 and 3-4 show frequency histograms of EBV accuracy based on linear and logistic animal models of fracture.

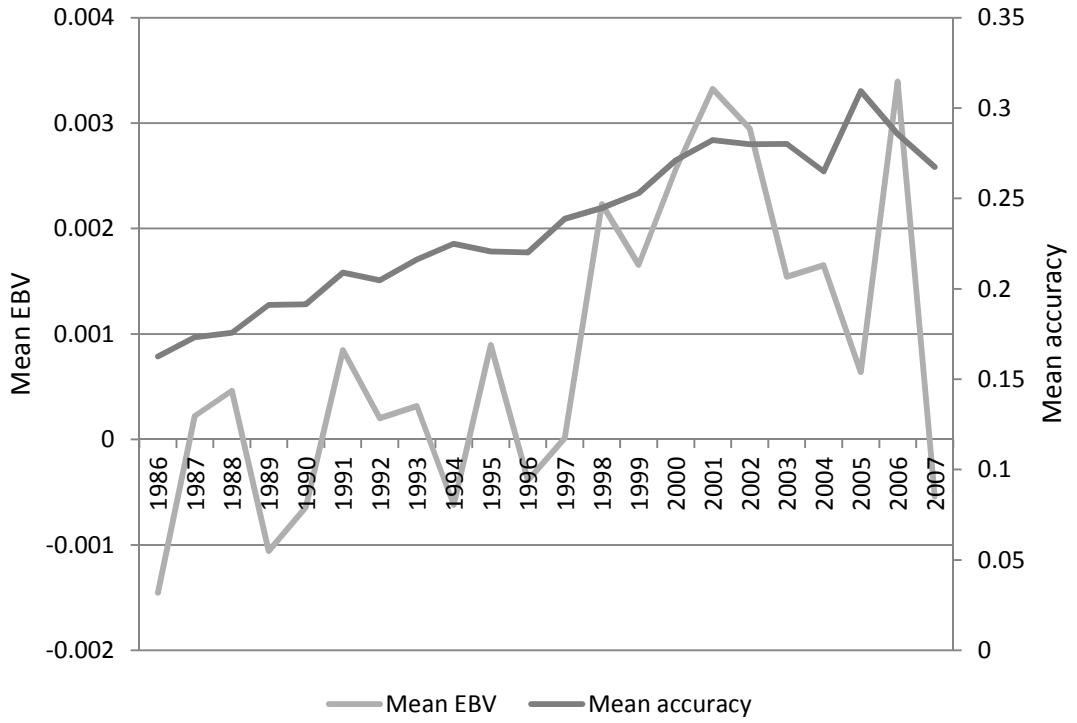

Figure 3-1. Mean EBV and mean EBV accuracies per year of birth cohort for fracture, derived from the final linear animal model.

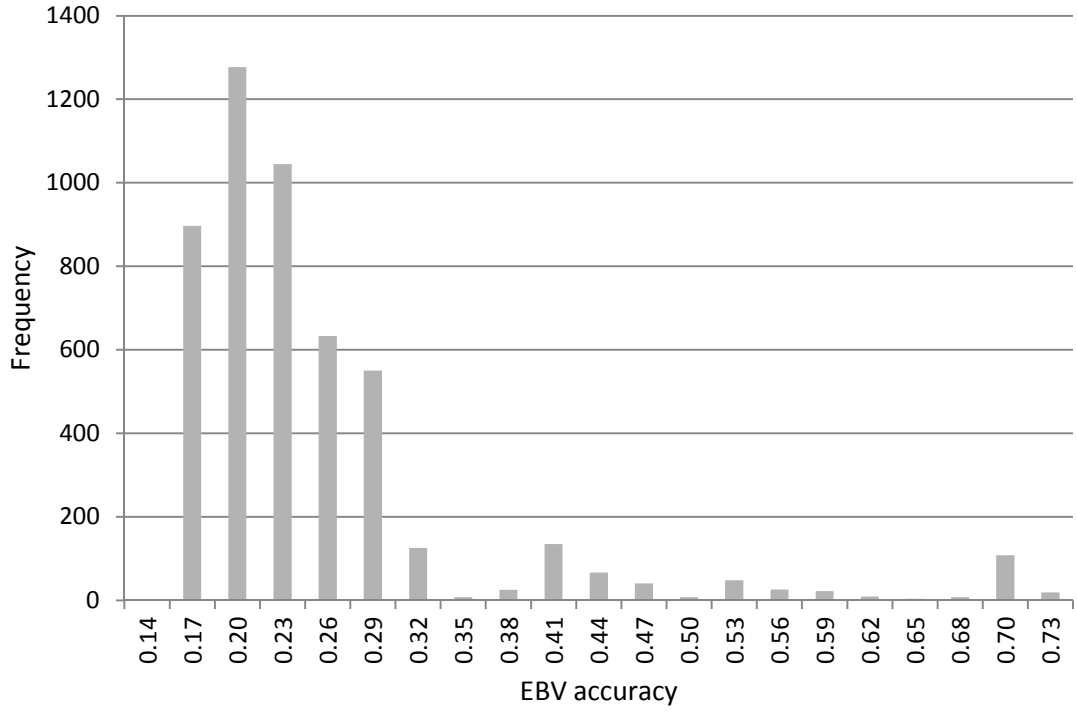

Figure 3-2. Frequency histogram of fracture EBV accuracies derived from the final linear animal model.

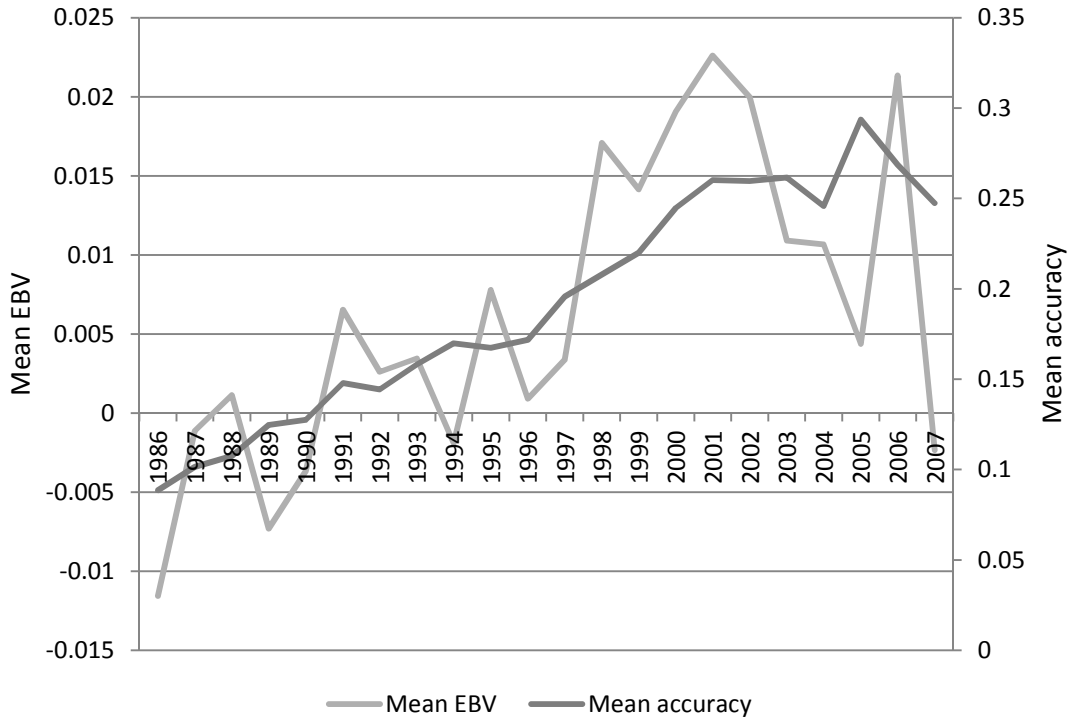

Figure 3-3. Mean EBV and mean EBV accuracies per year of birth cohort for fracture, derived from the final logistic animal model.

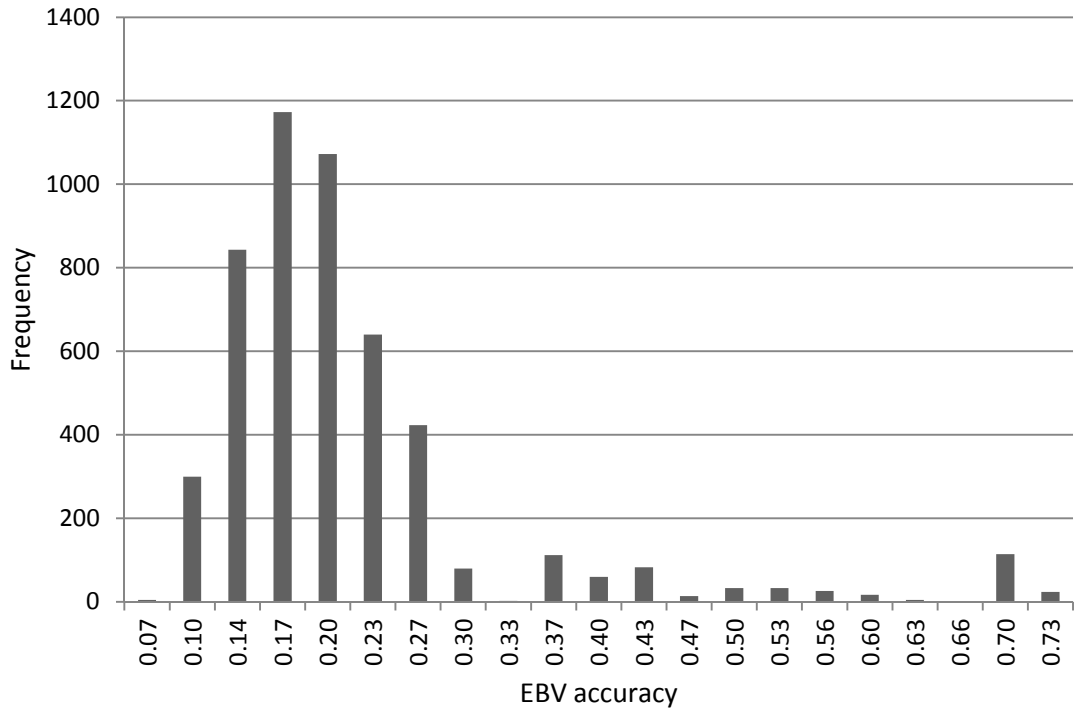

Figure 3-4. Frequency histogram of fracture EBV accuracies, derived from the final logistic animal model.

Regression of logistic EBVs on year of birth was equivalent to an increase of 0.0013 ± 0.0002 grades per annum, indicating a small but significant increasing (worsening) trend. Spearman rank correlation coefficient for EBVs generated using linear and logistic models

was 0.997 with a *p*-value <0.001, indicating close agreement in the ranking of animals based on either model.

3.1.4 Hong Kong fracture discussion

The prevalence of fracture in the population of horses in Hong Kong between September 1996 and December 2010 was 13.4%. Three percent of horses were diagnosed with a distal limb fracture over the same period. These figures are in agreement with previous studies of fracture prevalence (Ely et al., 2004; Verheyen and Wood, 2004). The most prevalent fracture sites diagnosed at OVEs were the diaphysis of the metatarsal or metacarpal bones, the carpus and proximal sesamoid bones, in descending order (Figure 2-4). However, in retirement records, the most prevalent fracture sites diagnosed were the proximal sesamoid bones, carpus and proximal phalanx (Figure 2-8). This suggests that diaphyseal fractures are not as commonly career-ending as proximal sesamoid bone fractures. The three most prevalent fracture types diagnosed at OVEs and the five most prevalent types diagnosed in retirement records were all fractures of the distal limb (inclusive of carpus/tarsus). This finding reiterates the importance of distal limb fractures to the overall fracture caseload in this population. Case numbers of specific fracture types (e.g. proximal sesamoid bone fracture, proximal phalanx fracture) were too few to enable modelling of heritability. Although the prevalence of distal limb fracture was high enough to enable modelling, there was insufficient evidence in the current dataset to warrant heritability analysis (Table 3-2). This does not exclude the possibility of a heritable component to distal limb fracture risk, but rather suggests that there were insufficient cases in the data to detect such an effect.

For binary data, heritability on the observed scale is an underestimate of heritability on the continuous liability scale (Gianola, 1982); however, both linear and nonlinear methods have been employed in the analysis of binary data (Vazquez et al., 2009; Pritchard et al., 2012). The fact that sire or animal random variables significantly improved the fit of both logistic and linear models of fracture can be taken as highly suggestive of significant heritability.

Only one explanatory variable was found to be associated with fracture risk. The odds of fracture increased as year of birth increased. It is possible that year of birth is partly

accounting for different genetic groups of horses, and thus effecting a change in the odds of fracture over time. Alternatively, changes to horse management, and sensitivity of veterinary diagnoses may have resulted in this finding. If true, this association indicates a worrying trend towards increasing fracture risk in the future, which is supported by the broadly increasing trend in genetic liability shown in Figures 3-1 and 3-3.

Heritability estimates of fracture ranged from 0.03 to 0.11 with large standard errors. It is possible that the heritability of specific fracture types exceeds these estimates, but it was not possible to model such specific outcomes in this study due to limited case numbers. Fracture was modelled as a binary trait, and the higher heritability estimates obtained with logistic models suggest that these are more appropriate than linear models in this instance, as they capture more of the genetic variance. Heritability estimates obtained using animal models of fracture were consistently lower than estimates obtained using sire models. This may have occurred if there were many full-sibling families in the pedigree, as the genetic covariance of full siblings is $\frac{1}{2}\sigma_a^2 + \frac{1}{4}\sigma_d^2$(where subscript *d* stands for dominance), which may have inflated the estimate of additive genetic variance for the sire model. Standard errors associated with animal model heritability estimates were smaller than those of sire models, suggesting that heritability of fracture can be estimated with greater confidence using the animal model approach (Weideman et al., 2004; Stock and Distl, 2006).

Three significant, positive genetic correlations were found between fracture and other musculoskeletal conditions. Sire model genetic correlations between fracture and ligament injury, and fracture and suspensory ligament injury were significantly greater than zero, and both animal and sire model genetic correlations between fracture and OA were significant. Positive genetic correlations arise between conditions if the genes coding for these conditions are the same (i.e. the genes are exhibiting pleiotropy), or are in linkage disequilibrium with each other. In an animal health context, positive genetic correlations between deleterious conditions can be useful, as selective breeding that reduces the incidence of one condition, may effect a reduction in the correlated condition concurrently. One study has linked injury of the suspensory ligament with lateral condylar fractures of the third metacarpal bone (Le Jeune et al., 2003). This study posited that damage to the medial branch of the suspensory ligament led to increased surface bone strains in areas that were subsequently more likely to fracture. This pathogenesis suggests an 'environmental' rather than genetic correlation between suspensory ligament injury and fracture, whereas in the current study, there was no significant phenotypic correlation between fracture and

conditions of the suspensory ligament. Fracture in this context encompassed all types and locations of fracture, thus relationships of subsets of fractures (e.g. lateral condyle) to suspensory ligament injury could not be ascertained.

OA has long been known as a possible sequel to articular fracture. These conditions were shown to be strongly genetically correlated in this study, suggesting pleiotropy of underlying risk alleles, or linkage between the genetic areas responsible for conferring risk of each condition. As well as this genetic relationship, it is also possible that the conditions may be environmentally linked, as destabilisation of a joint through articular fracture, with subsequent inflammation, could initiate inflammatory pathways leading to OA (provided the diagnoses of fracture and OA occurred in the same area of the same limb).

The genetic correlation, using an animal model, between all ligament injuries, and suspensory ligament injury was 0.99 (s.e. 0.03) (see Chapter 4), indicating that these conditions can be considered as the same genetic entity, despite the inclusion of injuries to alternative ligaments in the definition of 'ligament conditions'.

No significant genetic correlation was found between fracture and EIPH/epistaxis in this study. It is possible that, when called to examine a horse that has a suspected fracture, the attending veterinarian would be unlikely to record epistaxis concurrently, as the seriousness of a fracture would 'outweigh' that of a case of EIPH/epistaxis. This could account for a negative phenotypic correlation. This postulated hierarchy of 'importance' in recording of clinical problems may also explain the significant negative phenotypic correlation between fracture and tendon injury, or alternatively, clinical signs of one of these conditions could mask those of the other, if they occurred in the same limb (e.g. lameness, swelling, pain).

Despite the positive genetic correlation between fracture and suspensory ligament injury (sire model), these conditions were not phenotypically correlated. Despite pleiotropy or linkage of the causative genes, environmental influences may have had an opposing effect, thus no phenotypic correlation was found.

Figures 3-1 and 3-3 depict how mean EBV for fracture, per year of birth cohort, has changed over approximately twenty years. Both figures show an overall increasing trend

of EBVs, with many fluctuations. For a binary trait where 1 indicates a disease state and 0 indicates freedom from the disease, positive EBVs describe animals that are more likely to suffer the disease compared with their peers. This trend is worrisome, and the reasons behind this increase in risk are unknown. Mean EBV accuracies were disappointing at 0.24 for linear models and 0.20 for logistic models of fracture. Highly accurate EBVs allow faster responses to selection, thus breeding against fracture using the current estimates would offer only slow reductions in genetic risk. Mean accuracy of EBVs appeared to have increased over the study period in a consistent trend, thus continued selection based on fracture EBVs would enable acceleration in genetic gain. This increase in accuracy may be reflecting the greater amount of information available in the latter years of the study.

In conclusion, the heritability of fracture in the Hong Kong population appears to be low, but significant. Animal models on the logistic scale should be used to analyse fracture as a binary trait, as they take into account all relationships in a pedigree and give estimates with narrow standard errors. The genetic risk of fracture in the population increased over the study period, but EBV accuracies improved, thus selective breeding in the countries of origin would be timely and viable. Positive genetic correlations between fracture and OA, and between fracture and suspensory ligament injury, could potentially be utilised to reduce the risk of all of these conditions concurrently.

3.2 BHA fracture analysis results

Of 78151 horses in the BHA data, 797 horses were diagnosed with distal limb fractures, giving a crude prevalence of 1%. Over all race types, there were 0.89 fractures per 1000 race starts (799 in 900757 starts). Per 1000 race starts, the number of fractures recorded during flat, hurdle, NHF, and steeplechase races were 0.51, 1.44, 0.18 and 1.75, respectively.

Table 3-5 shows the results of a bivariate animal model analysis of flat and jump racing as dependent variables. In this analysis, a horse was classified as a flat racing horse if it competed in any flat races, and jump racing horses were those that competed in any jump races. Both heritability estimates are large and significant indicating the propensity for the progeny of flat racing horses to be involved in flat racing, and likewise for jump racing

horses' progeny. Due to the significant large negative genetic correlation between flat and jump racing, they were considered as separate populations, and distal limb fracture was analysed for the entire population, as well as the flat racing and jump racing populations separately.

Table 3-5. Estimated heritability of competing in flat racing and competing in jump racing, and the genetic correlation between them.

	Estimate	Standard error	95% CI
Flat racing h^2	0.9131	0.0054	0.9025 – 0.9237
Jump racing h^2	0.5442	0.0075	0.5397 – 0.5589
r_g* between jump and flat racing horses	-0.8569	0.0059	-0.8685 - -0.8453

*Genetic correlation.

3.2.1 Results from analysis using all horses

3.2.1.1 'Per Horse' (PH) dataset multivariable analysis

Table 3-6 gives the *p*-values of likelihood ratio tests performed on distal limb fracture final models, to ascertain whether addition of the genetic random effect variable improved the model. This was the case for all models except the logistic sire model, where addition of sire significantly worsened the log likelihood.

Table 3-6. Likelihood ratio test statistics and *p*-values of final models of distal limb fracture. LRT statistic is determined by $2(logL_{full} - logL_{red})$ where *L* is likelihood, and *full* and *red* indicate the final model, and the 'reduced' final model, i.e. minus the genetic random variable only (sire or animal).

Model form		LRT Statistic	*p*-value
Sire	Linear	22.38	<0.001
	Logistic	-9703.74	NA
Animal	Linear	24.06	<0.001
	Logistic	32764.62	<0.001

Two fixed effects were found to be associated with distal limb fracture (Table 3-7). Race profile describes whether a horse exclusively ran flat races (coded group 1), or whether it ran any jump races (coded group 2). Those horses in group 2 were up to twice as likely as group 1 to sustain a distal limb fracture. Increasing number of starts was associated with

increased odds of distal limb fracture. Logistic model results suggest that for every 100 additional starts a horse made over its career, the odds of distal limb fracture doubled.

Table 3-7. Results of multivariable sire regression models on linear and logistic scales. Two fixed effects were significant in the final models; racing profile of the horse, and number of starts in the career.

		Linear model	Logistic model		
		Wald *p*-value	**Wald *p*-value**	**Odds ratio**	**95% CI**
Race profile	*1**	<0.001	<0.001	1 (REF)	
	2†			2.008	1.705 – 2.366
Number of career starts (hundreds)		<0.001	<0.001	2.123	1.348 – 3.344

* Horses that competed exclusively in flat racing; † horses that competed in any jump racing.

3.2.1.2 PH dataset heritability and genetic correlation estimation

Table 3-8 gives the heritability estimates for distal limb fracture based on sire and animal linear and logistic models. All estimates were significantly greater than zero. Logistic heritability estimates were 15 to 22 times larger than linear estimates. Confidence intervals for linear model estimates overlap each other, as do confidence intervals for logistic models. Animal and sire model estimates were similar to each other in both linear and logistic models.

Table 3-8. Heritability estimates of distal limb fracture in the BHA PH dataset.

Model form		h^2	**Standard error**	**95% CI***
Sire	Linear	0.0096	0.0026	0.0045 – 0.0192
	Logistic	0.2161	0.0637	0.0536 – 0.3786
Animal	Linear	0.0102	0.0026	0.0051 – 0.0204
	Logistic	0.1553	0.0462	0.0248 – 0.1801

*Confidence interval (1.96 x standard error)

Genetic and phenotypic correlations between distal limb fracture, epistaxis and SDFT injury in the PH BHA dataset are shown in Table 3-9. Distal limb fracture and SDFT injury were found to be significantly genetically correlated, with a correlation coefficient of 0.43 to 0.47. No significant genetic correlation was found between epistaxis and distal

limb fracture. There were no statistically significant phenotypic correlations found between any pair of conditions.

Table 3-9. Genetic and phenotypic correlations (with standard errors and 95% confidence intervals) between fracture , epistaxis and SDFT injury. Sire model results appear above animal model results, which are in bold type. P-values are based on LRT between the unconstrained model and the same model with the genetic correlation constrained to 0.00001.

	r_g*	s.e.†	95% CI‡	*p*- value	r_p§	s.e.	95% CI
Epistaxis	0.1985 **0.1753**	0.1221 **0.1166**	-0.0408 – 0.4378 **-0.0532 – 0.4638**	0.71 **0.24**	0.0004 **0.0003**	0.0036 **0.0036**	-0.0067 – 0.0075 **-0.0068 – 0.0074**
SDFT injury	0.4661 **0.4268**	0.0996 **0.0961**	0.2509 – 0.6613 **0.2385 – 0.6152**	<0.001 **<0.001**	-0.0140 **-0.0141**	0.0036 **0.0036**	-0.0211 – 0.0069 **-0.0212 – 0.0070**

*Genetic correlation, †standard error, ‡confidence interval, §phenotypic correlation.

3.2.1.3 PH dataset estimated breeding values for distal limb fracture

Figures 3-5 and 3-7 show the mean EBV for distal limb fracture, and the mean accuracy of EBVs per year of birth cohort. Both appear to show that distal limb fracture EBVs reduced (improved) over the study period, becoming negative, on average, around the year 2000. Regression of logistic EBVs on year of birth was equivalent to a decrease of -0.0049 ± 0.0001 grades per annum, indicating a small but significant decreasing (improving) trend. Mean EBV accuracy was 0.21 for logistic models (range 0.06 – 0.59) and 0.26 for linear models (range 0.10 – 0.70) (Figures 3-6 and 3-8). The mean EBV accuracy per year of birth cohort appeared to increase over the study period. Spearman rank correlation coefficient for EBVs generated using linear and logistic animal models was 0.997 (*p*-value <0.001).

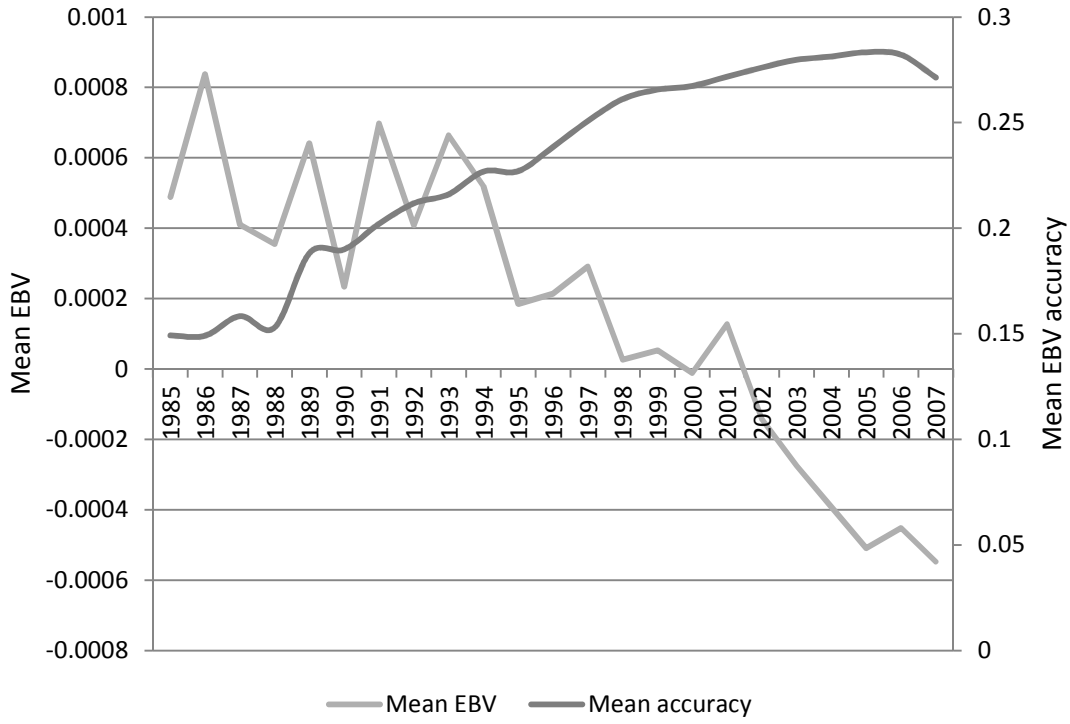

Figure 3-5. Mean EBV and mean EBV accuracies per year of birth cohort for distal limb fracture, derived from the final linear animal model.

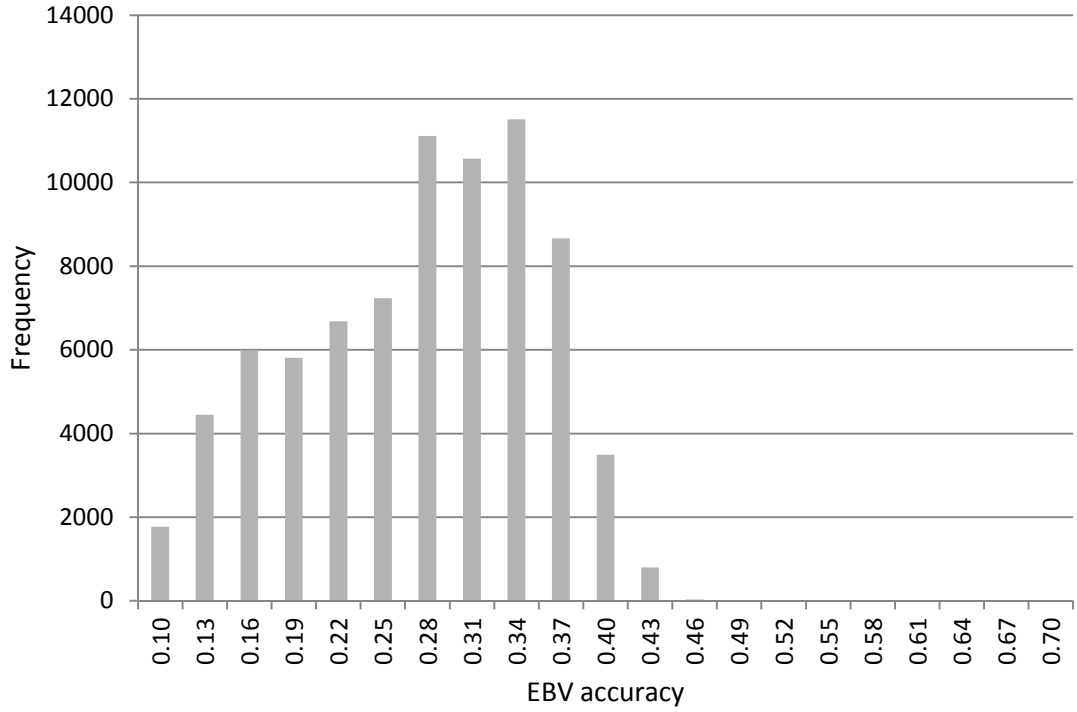

Figure 3-6. Frequency histogram of the accuracy of EBVs for distal limb fracture derived from the final linear animal model.

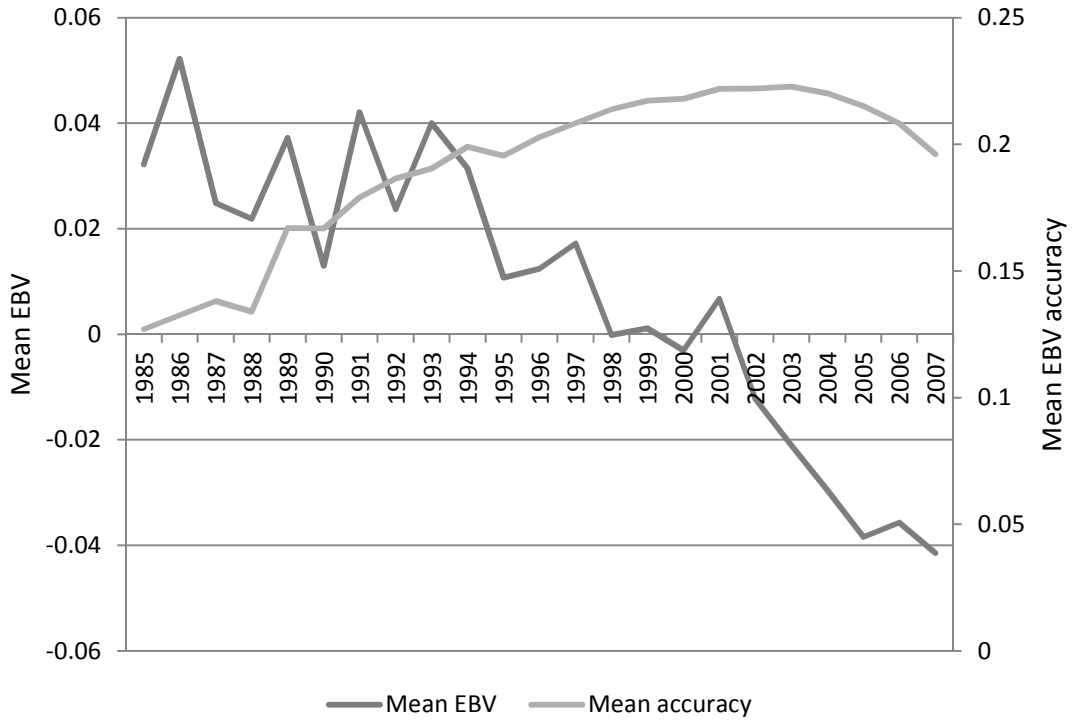

Figure 3-7. Mean EBV and mean EBV accuracies per year of birth cohort for distal limb fracture, derived from the final logistic animal model.

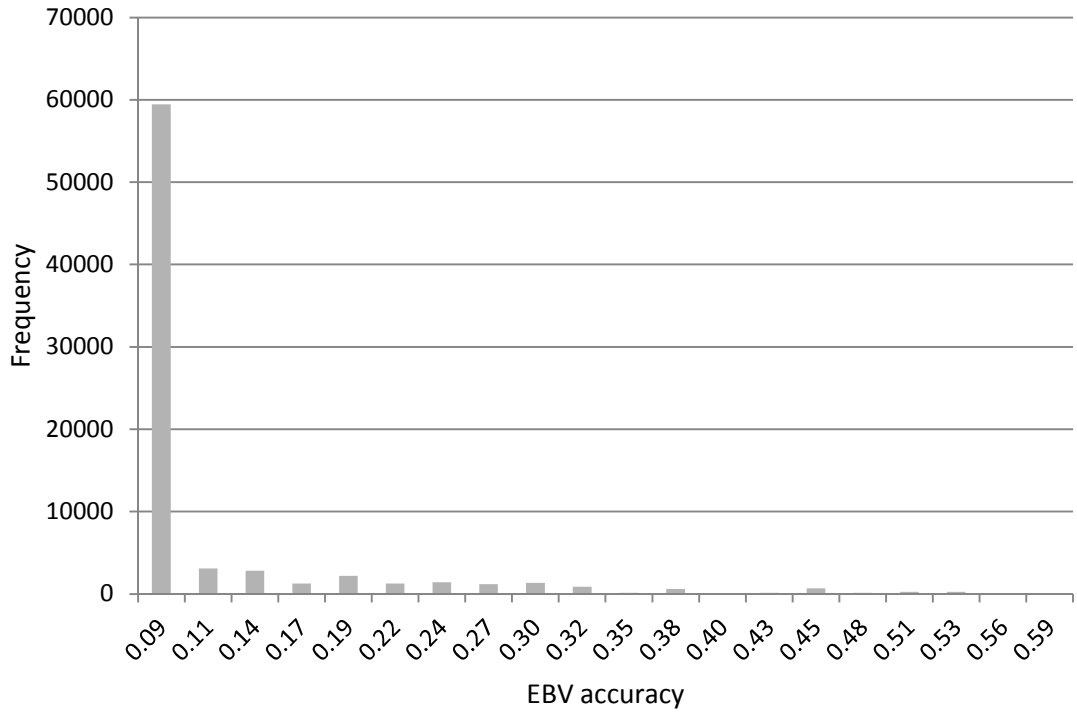

Figure 3-8. Frequency histogram of the accuracy of EBVs for distal limb fracture derived from the final logistic animal model.

3.2.1.4 BHA Repeated Records (REP) dataset heritability and repeatability

Heritability, repeatability and permanent environmental variance estimates are shown in Table 3-10. Sire and animal model heritability estimates were similar to each other. Permanent environmental variance was negligible in both models, and repeatability was similar to heritability in the animal model.

Table 3-10. Heritability, repeatabilty and permanent environmental variance of distal limb fracture, based on linear repeatability models.

	h^2* (s.e.)	σ^2_{pe}† (s.e.)	r‡
Sire model	0.0010 (0.0002)	<0.0000 (<0.0000)	0.0002 (0.0001)
Animal model	0.0009 (0.0002)	<0.0000 (<0.0000)	0.0009 (0.0002)

*Heritability;† permanent environmental variance;‡ repeatability.

3.2.2 Flat racing population model results

3.2.2.1 Multivariable analysis

Two fixed effects were found to be associated with distal limb fracture in the flat racing Thoroughbred population. There were 49660 horses that competed in any flat race over the study period, of which 56.7% (28142) were male. These horses were diagnosed with 419 distal limb fractures at the racecourse, 317 (75.7%) of which were diagnosed in males and 102 (24.3%) in females. Female sex was associated with decreased odds of distal limb fracture compared to male sex, and increased career length in years was associated with increased odds of distal limb fracture (Table 3-11).

Table 3-11. Results of multivariable sire regression models on linear and logistic scales. Two fixed effects were significant in the final models: sex and career length in years.

		Linear model	Logistic model		
		Wald *p*-value	Wald *p*-value	Odds ratio	95% CI*
Sex	Male	<0.001	<0.001	1 (REF)	
	Female			0.454	0.357-0.577
Career length (years)		<0.001	0.002	1.085	1.029-1.145

*Confidence interval (1.96 x standard error)

In all models of distal limb fracture in flat racing horses, inclusion of a sire or animal random effect significantly improved the fit of the model (*p*-value <0.05) (Table 3-12).

Table 3-12. Likelihood ratio test statistics and *p*-values of final models in the flat racing horse population of distal limb fracture. LRT statistic is determined by $2(logL_{full} - logL_{red})$ where L is likelihood, and *full* and *red* indicate the final model, and the 'reduced' final model, i.e. minus the genetic random variable only (sire or animal).

Model form		LRT statistic	*p*-value
Sire	Linear	7.60	0.006
	Logistic	6599.40	<0.001
Animal	Linear	8.16	0.004
	Logistic	20169.42	<0.001

Heritability estimates of distal limb fracture at the racecourse among flat racing horses ranged from 0.01 to 0.20 (Table 3-13). All estimates were significantly greater than zero. Logistic estimates were 20 to 28 times larger than estimates on the observed binary (linear model) scale. Heritability estimates using sire models were similar to estimates using animal models. Standard errors of heritability estimates from linear models were very similar, whereas the logistic animal model standard error was smaller than the logistic sire model standard error.

Table 3-13. Heritability of distal limb fracture within the flat racing population in the PH BHA dataset.

Model form		h^2	Standard error	95% CI*
Sire	Linear	0.0071	0.0030	0.0012 - 0.0130
	Logistic	0.2024	0.0920	0.0221 – 0.3826
Animal	Linear	0.0073	0.0029	0.0016 – 0.0130
	Logistic	0.1464	0.0678	0.0136 – 0.2792

*Confidence interval (1.96 x standard error)

3.2.3 Jump racing population model results

3.2.3.1 Multivariable analysis

Table 3-14 shows the results of multivariable modelling of distal limb fracture among horses that competed in jump racing during the study period. Of 40135 horses that competed in jump racing over the study period, 30464 (75.9%) were male. These horses were diagnosed with 545 distal limb fractures, 449 (82.4%) of which occurred in males and 96 (17.6%) in females. Only sex was found to be associated with distal limb fracture. Female sex was associated with reduced odds of distal limb fracture.

Table 3-14. Results of multivariable sire regression models on linear and logistic scales. Only sex was significant in the final model.

		Linear model	**Logistic model**		
		Wald *p*-value	**Wald *p*-value**	**Odds ratio**	**95% CI***
Sex	Male	<0.001	<0.001	1 (REF)	
	Female			0.681	0.539 – 0.860

*Confidence interval (1.96 x standard error)

The inclusion of animal or sire as a random genetic variable significantly improved the fit of every model of distal limb fracture in jump racing horses (Table 3-15).

Table 3-15. Likelihood ratio test statistics and *p*-values of final models in the jump racing horse population of distal limb fracture. LRT statistic is determined by $2(logL_{full} - logL_{red})$ where L is likelihood, and *full* and *red* indicate the final model, and the 'reduced' final model, i.e. minus the genetic random variable only (sire or animal).

Model form		**LRT statistic**	***p*-value**
Sire	Linear	14.72	<0.001
	Logistic	5850.70	<0.001
Animal	Linear	15.48	<0.001
	Logistic	16960.72	<0.001

Heritability estimates of distal limb fracture at the racecourse among jump racing horses ranged from 0.01 to 0.24 (Table 3-16). All estimates were significantly greater than zero. Estimates made on the logistic scale were 12 to 18 times larger than those on the observed binary scale. Linear model estimates from sire and animal models were similar, as were logistic model estimates. Linear model standard errors were identical to four decimal

places, but the logistic animal model standard error was smaller than the standard error from the logistic sire model.

Table 3-16. Heritability of distal limb fracture within the jump racing population in the PH BHA dataset.

Model form		h^2	Standard error	95% CI*
Sire	Linear	0.0129	0.0044	0.0043 – 0.0215
	Logistic	0.2390	0.0845	0.0735 – 0.4045
Animal	Linear	0.0132	0.0044	0.0046 – 0.0218
	Logistic	0.1583	0.0574	0.0459 – 0.2707

*Confidence interval (1.96 x standard error)

3.2.4 BHA distal limb fracture discussion

The prevalence of distal limb fracture was similar to previous studies, with the exception of NHF distal limb fracture prevalence, which was lower in the current study compared with Parkin et al. (2004) (Parkin et al., 2004b; Pinchbeck, 2004). Fracture prevalence was greater in jump racing compared to flat racing, as has been found in previous studies (Parkin et al., 2004b; Kristoffersen et al., 2010b). This relationship was also observed in the multivariable modelling section, where the 'race profile' of a horse was associated with the risk of distal limb fracture. Jump racing places extra strain on musculoskeletal tissues, especially upon landing from fences. It is thought that these strains are an important contributor to increased fracture risk, along with a higher risk of falls and collisions during racing (Pinchbeck, 2004; Boden et al., 2005). This study showed that horses that compete in flat racing are genetically distinct, based on a significant negative genetic correlation, from those that compete in jump racing. For this reason, the heritability of distal limb fracture was also modelled in these populations separately.

The number of starts was found to be associated with increased odds of distal limb fracture when all horses were considered together. It is hypothesised that most distal limb fractures are fatigue fractures, therefore running more races represents accumulation of more microdamage, possibly leading to fracture. This variable may also be more simply explained as a measure of increased time at risk. To ascertain which of these explanations is correct, histological studies would be needed to examine the relationship of microdamage with the number of starts . This was beyond the scope of this study, but

accumulation of microdamage and functional adaptation of bone to exercise have been demonstrated extensively in literature (Lanyon, 1987; Nunamaker et al., 1990; Reilly et al., 1997; Boyde and Firth, 2005; Entwistle et al., 2008).

The heritability estimates of distal limb fracture in this study were significantly greater than zero, which, combined with the LRT *p*-values in Table 3-5, provide strong evidence that distal limb fracture is a heritable condition in this Thoroughbred population. Based on linear models, distal limb heritability was estimated at around 0.01, and logistic model heritability estimates were large at 0.16 and 0.22 in animal and sire models, respectively. Animal models on the logistic scale gave large heritability estimates with comparatively narrow confidence intervals, and produced EBVs that were ranked very similarly to those derived from linear models, thus they are suggested as the most appropriate model type for this condition in this population. It should be remembered that the dependent variable in this section of the BHA data is distal limb fracture at the racecourse, and does not include any fractures sustained or diagnosed away from the racetrack. The constraint of the case definition to include only fractures at the racecourse may have led to many 'heritable' fractures sustained during training being missed, with the effect of an artificial reduction in heritability estimates.

Distal limb fracture was found to be positively genetically correlated with injury to the SDFT. This finding suggests that a proportion of the genes giving rise to the risks of these conditions exhibit pleiotropy, or are in linkage disequilibrium with each other. Molecular genetic studies would be needed to ascertain which genes of importance for distal limb fracture and SDFT injury are related in this way. The sign of the correlation indicates that a favourable response in one of the conditions may occur when the other condition is targeted by selective breeding.

Estimated breeding values for distal limb fracture showed an improving trend over the study period, becoming negative (i.e. reduced risk relative to the population) around the year 2000. It is unknown why this trend was occurring. EBV accuracies also improved over the same period, meaning that the response to selective breeding based on these EBVs would have accelerated over time. EBV accuracy is calculated as follows:

$$accuracy = \sqrt{1 - \left(\frac{s.e.^2}{(1 + f_i)\sigma_a^2}\right)}$$

where *s.e.* is the standard error of the best linear unbiased prediction ('BLUP') breeding value ('BV') (square root of prediction error variance) produced by ASReml, and *f* is the inbreeding coefficient of animal *i*. Accuracy therefore increases with increasing inbreeding, increasing additive genetic variance, or decreasing BLUP BV standard error. Average inbreeding was negligible in all but one year of birth cohort (information not shown), therefore it is thought that a change in inbreeding coefficient is not responsible for the changing EBV accuracy. The mean standard error (produced by ASReml) of the estimated breeding values per year of birth cohort is shown in Figure 3-9.

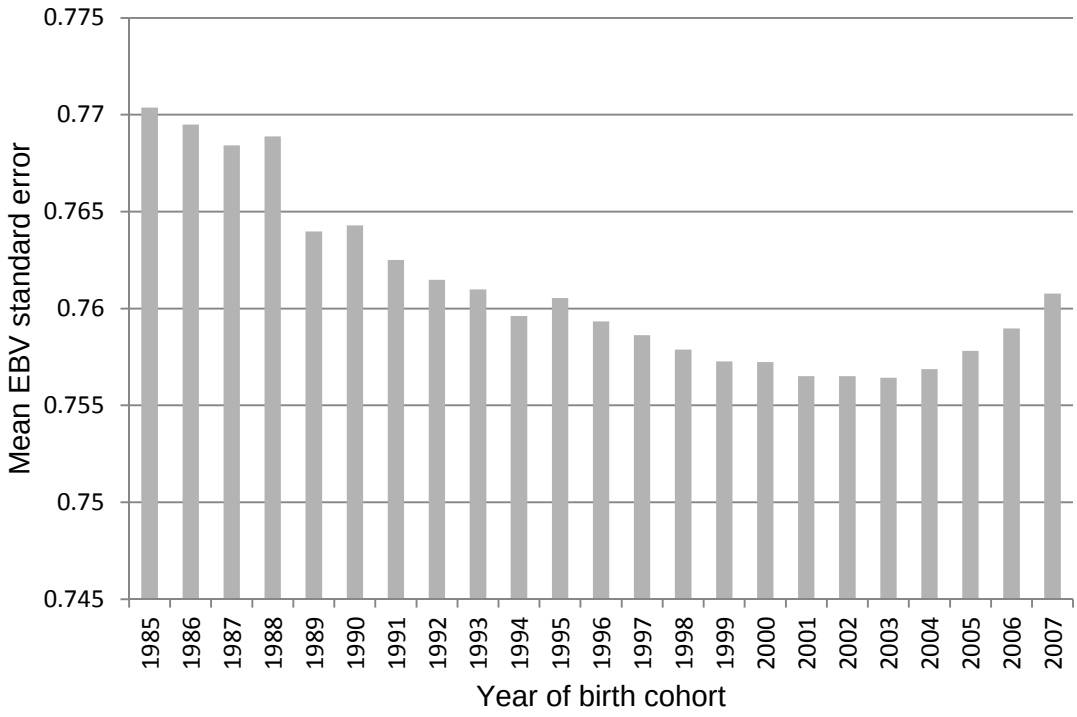

Figure 3-9. Mean standard error of distal limb fracture EBV (logistic model) per year of birth cohort.

The number of horses that contributed phenotypic information to the BHA data per year of birth cohort showed the same trend as Figure 3-9 inversed (Figure 2.9). The greater the amount of information on which the BVs are based, the more accurate they will be, therefore, the smaller the standard errors of the breeding values, and the greater the EBV accuracy. To ascertain whether alterations occurred in the genetic variance for distal limb fracture over the study period, the data were split into year of birth cohort groups and animal models of distal limb fracture were run using the same model form as used previously. The genetic variance is shown below (Table 3-17).

Table 3-17. Genetic variance of distal limb fracture over the study period. Genetic variance was calculated by ASReml based on identical logistic animal models of distal limb fracture, with fixed effects of racing profile and number of starts included.

Year of birth cohort group	Number of horses in group	Distal limb fracture genetic variance
1983-1989	713	n/a*
1990-1994	7515	0.8519
1995-1999	24102	0.6729
2000-2004	32372	0.5682
2005-2007	13449	1.4481

*Non-convergence of model.

The values of genetic variance in Table 3-17 demonstrate that mean genetic variance is a function of the amount of data available, thus the more information included in the calculation, the smaller the mean variance as it approaches the true population mean. Thus, the changing EBV accuracy in the BHA data could be due to altered genetic variance, BLUP EBV standard error, and/or the number of horses per year of birth cohort, which affects both.

Repeatability of distal limb fracture was low to moderate, and was accounted for almost entirely by the heritability, with negligible permanent environmental effects detected. Repeatability may be artificially low in this study as distal limb fracture was recorded as a binary trait, rather than a continuous or polychotomous variable. These results indicate that genetic influences are more important than permanent environmental effects in the risk of distal limb fracture.

As mentioned above, flat and jump racing populations were also considered separately in this part of the study. Distal limb fracture was significantly heritable in all models of both flat and jump racing populations. Whereas racing profile and number of starts were associated with fracture risk when all horses were considered together, significant fixed effects in flat and jump racing populations differed. In flat racing, sex and career length were associated with distal limb fracture, and in jump racing, only sex was found to be associated. The effect of sex on the risk of distal limb fracture in the population as a whole may have been masked by the strong relationship found between race profile and distal limb fracture, where jump racing was associated with increased odds of up to 2.4 times. Female sex was associated with reduced odds of fracture in both flat and jump racing. This could be a sex hormone effect, but the true reason behind this association remains

unknown. Previous studies have detected an association between sex and serious injury, and males have been shown in these studies to be at increased risk (Estberg et al., 1996; Estberg et al., 1998; Cohen et al., 2000; Hernandez et al., 2001).

Horses that had longer flat racing careers were at increased risk of distal limb fracture, but this association was not seen in jump racing. The average length of racing career for jump racing horses in this study was 1.9 years, and for flat racing the average career length was 1.5 years. Over this time, flat racing horses, on average, competed in 13.5 races, whereas jump racing horses competed in only 12.2 races. The association found between career length and distal limb fracture in flat racing, but not jump racing could reflect the shorter careers of flat racing horses, in which more races are run compared with jump racing. This intensity and duration of racing (and presumably training) activity could have led to faster accumulation of microdamage in skeletal tissues, with an ensuing risk of fracture.

The prevalence of distal limb fracture on the racecourse in this study was higher in jump racing compared to flat racing. More falls and collisions occur in jump racing compared with racing on the flat, and all fracture types were included in the case definition for this study. Because falls and collisions, and therefore also the monotonic fractures they cause, can occur at any time in a racing career, any association between accumulated exercise and increased fracture risk could have been masked in the BHA dataset.

The heritability of distal limb fracture was estimated at 0.01 using the observed binary scale (linear models). Logistic model estimates were around 0.15 to 0.24 for models where all horses were considered together, and also in flat and jump racing horses. All confidence intervals for linear model heritability estimates overlapped, suggesting that they were not significantly different from each other. Similarly, all confidence intervals for logistic model estimates overlapped.

3.3 General fracture discussion

At the time of writing, no studies have been published that estimate the heritability of fracture in the racehorse. This study provides estimates based on two separate populations of Thoroughbreds, where fracture was recorded at different times (in racing, or in racing and training), and used both linear and logistic animal and sire models. In both datasets,

fracture (or distal limb fracture) prevalence was similar to previous studies. Differences between the two datasets will be discussed further in a later general discussion.

Racing over jumps has been shown to be a risk factor for serious musculoskeletal injuries in past studies (Bailey et al., 1998; Parkin et al., 2004a; Boden et al., 2005). Analysis of distal limb fracture on British racecourses in the present study also showed an increased risk in steeplechase or hurdle races compared with flat races. This study has shown that the populations of racehorses that compete in flat and jump races are genetically distinct, evidenced by the significant negative genetic correlation between these groups (Table 3-9). Distal limb fracture heritability was similar in both jump and flat racing. The greater prevalence of distal limb fracture in jump racing compared to flat racing may therefore be due to increased numbers of 'non-heritable' fractures in jump racing, i.e. monotonic fractures sustained during falls or collisions. Within jump racing, the prevalence of distal limb fracture was greater in steeplechase compared with hurdle races. Height, style and placement of jumps vary between racecourses and race types, which may account for the difference in fracture prevalence.

Heritability of fracture was estimated in the HK data to be between 0.03 and 0.11. All fractures were included in this analysis, and diagnoses could happen during racing or training activities. BHA heritability estimates of distal limb fracture during racing were consistently close to 0.01 using linear models, and ranged from 0.15 to 0.24 using logistic models. Both the BHA and HK data show evidence of the significance of fracture heritability, despite the different approaches used. Higher heritability estimates from the BHA data compared to the HK data may reflect their different case definitions (fracture versus distal limb fracture), when records were collected (at the racecourse or at any time), or may be due to the genetic differences between the horse populations in question. The magnitude of the heritability in both datasets is not negligible, and suggests that genetic risk could be amenable to change following targeted selective breeding.

The current study is not designed to investigate how the genes underlying fracture are involved in its aetiopathogenesis. The heritability of fracture could be explained by at least two different theories. The genes responsible may encode an anatomical weakness in specific sites that render the bone vulnerable to complete failure, which may be more likely during racing, training and particularly jumping, compared with resting. Alternatively, the underlying genetic risk could be due to a defect in the normal response of musculoskeletal

tissues to exercise. Should the biochemical and cellular processes that underpin healing of microdamaged bone be deranged, those sites that are under greatest strains would be the most likely to fail. Fractures of the distal limb predominate in most racing jurisdictions, and a large proportion of these are fatigue fractures, therefore it is postulated by the author that one or more fatigue fracture sites on the distal limb are responsible for the heritable nature of fracture reported here. This may explain the higher heritability estimates of distal limb fracture in the BHA data compared with all fractures in the HK data.

In order to determine the underlying pathogenesis of genetically-mediated fractures, further large studies of more specific fracture types are warranted. Both anatomical specificity and histopathological diagnoses would be needed in order to identify the precise definition of a 'heritable fracture'. The EBVs generated from these future studies would be likely to be more accurate, and therefore selective breeding away from heritable fracture types would be more realistic than attempting to breed away from all fracture types at once. Alternatively, once the characteristics of heritable fractures are identified, genomic selection (GS) could be implemented to produce genomic estimated breeding values (GEBVs) for fracture, which would further enhance the accuracy of selection, at the same time as reducing the generation interval, leading to an increased response to selection. If GS were to be attempted, a suitable Thoroughbred population of sufficient size (likely over 1000 animals) would need to be identified. This population would need to contain animals from the same genetic lines as the young stock to be used for breeding, in order that the SNPs are in LD with the relevant QTL. In a practical sense, for Hong Kong this would mean that the reference population must include lines from all the exporting countries from which Hong Kong accepts racehorses. All of this initial population would need to be genotyped and phenotyped for the heritable fracture phenotype, in order that the effects of the SNPs involved could be estimated. Young horses could then be genotyped and their GEBV ascertained from the summation of the effects of the SNPs they carry. The publication of mare and stallion GEBVs for fracture could be used to allow breeders to make informed decisions and avoid producing foals at high risk of fracture.

Another possibility would be to conduct studies to ascertain the genes responsible for a large proportion of fracture risk. Commercial tests could be offered that genotype horses at those important sites. This is risky, however, as interpretation of these tests in the context of an individual's risk is difficult. The identification of one or more risk alleles in an individual's genome does not equate to any level of certainty that that individual will

sustain the injury in question, as many other genes are likely to influence the injury risk, as well as a plethora of environmental influences. Breeders would require a large amount of guidance from the test providers and veterinarians, and would need to be counselled that a horse carrying risk alleles may never fracture, whereas one with no risk alleles is still susceptible to fracture.

Regardless of which approach is taken in future to minimise genetic risk, environmental risk will always remain significant in disease pathogenesis. Further high quality studies of environmental risk factors for fracture would be of use, but more importantly, the prominent findings of these studies must be communicated clearly to owners and trainers. Identification of modifiable risk factors is of no use unless they can affect a change in the training or management of racehorses to minimise future fracture risk. For this reason, validation of the useful findings of previous studies is essential, as is ensuring that communication within the racing industry allows dissemination of coordinated advice to all stakeholders to maintain the integrity of this research in the public sphere.

4. CHAPTER IV. MUSCULOSKELETAL CONDITIONS

4.1 Musculoskeletal disorders in Hong Kong

4.1.1 Ligament injuries

Both ligament injuries as a whole and suspensory ligament injuries specifically were modelled in this study. A total of 585 (11.6%) horses were diagnosed with at least one ligament injury over the study period. There were 61 (1.2%) horses in which the ligament involved was not the suspensory ligament (e.g. check ligament, sesamoidean ligament), or the ligament involved was not named. Thus 524 (10.4%) horses were diagnosed with a suspensory ligament injury. As mentioned previously, the genetic correlation found between ligament and suspensory ligament injury, using a linear animal model without fixed effects was 0.986 (s.e. 0.027), which suggests that they can be considered the same genetic trait. For this reason, only results of modelling suspensory ligament injury will be presented.

4.1.1.1 Results of multivariable modelling of suspensory ligament injury

Table 4-1 shows the LRT statistics of final multivariable models of suspensory ligament injury in the Hong Kong Thoroughbred population. All models were significantly improved by addition of the genetic random variable (*p*-value <0.05).

Table 4-1. Likelihood ratio test statistics and *p*-values of final models of suspensory ligament injury. LRT statistic is determined by $2(logL_{full} - logL_{red})$ where *L* is likelihood, and *full* and *red* indicate the final model, and the 'reduced' final model, i.e. minus the genetic random variable only (sire or animal).

Model form		LRT statistic	*p*-value
Sire	Linear	5.90	0.015
	Logistic	351.92	<0.001
Animal	Linear	5.72	0.017
	Logistic	630.46	<0.001

Three fixed effects were found to be significantly associated with suspensory ligament injury. The results of these models are shown in Table 4-2. The odds of suspensory ligament injury were found to increase with increasing age at retirement. Increasing number of starts was associated with reduced odds, and originating from Australasia or

North America compared to Europe was associated with reduced odds of suspensory ligament injury.

Table 4-2. Results of multivariable sire regression models of suspensory ligament injury on linear and logistic scales.

		Linear model	Logistic model		
		Wald *p*-value	**Wald *p*-value**	**Odds ratio**	**95% CI**
Age at retirement (yrs)		<0.001	<0.001	1.454	1.344 – 1.573
Number of starts		<0.001	<0.001	0.986	0.980 – 993
Continent group	*1**	<0.001	<0.001	1 (REF)	
	2†			0.707	0.554 – 0.902
	3‡			0.607	0.432 – 0.852
	4§			0.553	0.230 – 1.332

*Europe; †Austalasia; ‡North America; §other.

Table 4-3 shows the results of heritability analyses of suspensory ligament injury in the Hong Kong study population. All estimates exceeded their standard errors, but some confidence intervals spanned zero. Logistic estimates were two to three times larger than estimates made on the observed binary scale (linear models). Sire model estimates exceeded animal model estimates.

Table 4-3. Heritability estimates of suspensory ligament injury.

Model form		h^2	**Standard error**	**95% CI***
Sire	Linear	0.0602	0.0303	0.0008 – 0.1196
	Logistic	0.1703	0.0906	-0.0073 – 0.3479
Animal	Linear	0.0471	0.0239	0.0003 – 0.0939
	Logistic	0.0683	0.0424	-0.0148 – 0.1514

*Confidence interval (1.96 x standard error)

Genetic and phenotypic correlations between suspensory ligament injury and the other conditions studied in the HKJC data are shown in Table 4-4. Only suspensory ligament injury and fracture were (positively) genetically correlated (sire model). Suspensory ligament injury was negatively phenotypically correlated with EIPH/epistaxis and tendon injury, and positively phenotypically correlated with OA.

Table 4-4. Genetic and phenotypic correlations (with standard errors and 95% confidence intervals) between suspensory ligament injury and the named conditions. Sire model results appear above animal model results, which are in bold type. P-values are based on LRT between the unconstrained model and the same model with the genetic correlation constrained to 0.0001.

	r_g*	s.e.†	95% CI‡	*p*-value	r_p§	s.e.	95% CI
EIPH	-0.0139 **-0.3382**	0.2470 **0.2406**	-0.4980 – 0.4702 **-0.8098 – 0.1334**	1.0000 **0.1380**	-0.0301 **-0.0299**	0.0141 **0.0141**	-0.0577 - -0.0025 **-0.0575 - -0.0023**
Fracture	0.5902 **0.4225**	0.2941 **0.2912**	0.0138 – 1.1666 **-0.1483 – 0.9933**	0.0455 **0.1435**	-0.0051 **-0.0049**	0.0141 **0.0141**	-0.0327 – 0.0225 **-0.0325 – 0.0223**
OA	0.1162 **0.1326**	0.2426 **0.2389**	-0.3593 – 0.5917 **-0.3356 – 0.6008**	0.6390 **0.5716**	0.0333 **0.0339**	0.0141 **0.0141**	0.0057 – 0.0609 **0.0062 – 0.0615**
Tendon injury	-0.1302 **-0.1217**	0.2914 **0.2645**	-0.7013 – 0.4409 **-0.6401 – 0.3967**	0.6547 **0.6547**	-0.0636 **-0.0637**	0.0140 **0.0141**	-0.0914 - -0.0362 **-0.0913 – -0.0361**

*Genetic correlation, †standard error, ‡confidence interval, §phenotypic correlation.

EBVs were extracted from final animal model analyses, and their accuracies calculated. The EBV accuracy in linear models ranged from 0.20 to 0.75 (mean 0.31), and logistic models ranged from 0.08 to 0.74 (mean 0.25). Mean EBV and accuracy for each year of birth cohort based on linear and logistic models are shown in Figures 4-1 and 4-3. Figures 4-2 and 4-4 show frequency histograms of the EBV accuracies from linear and logistic models, respectively.

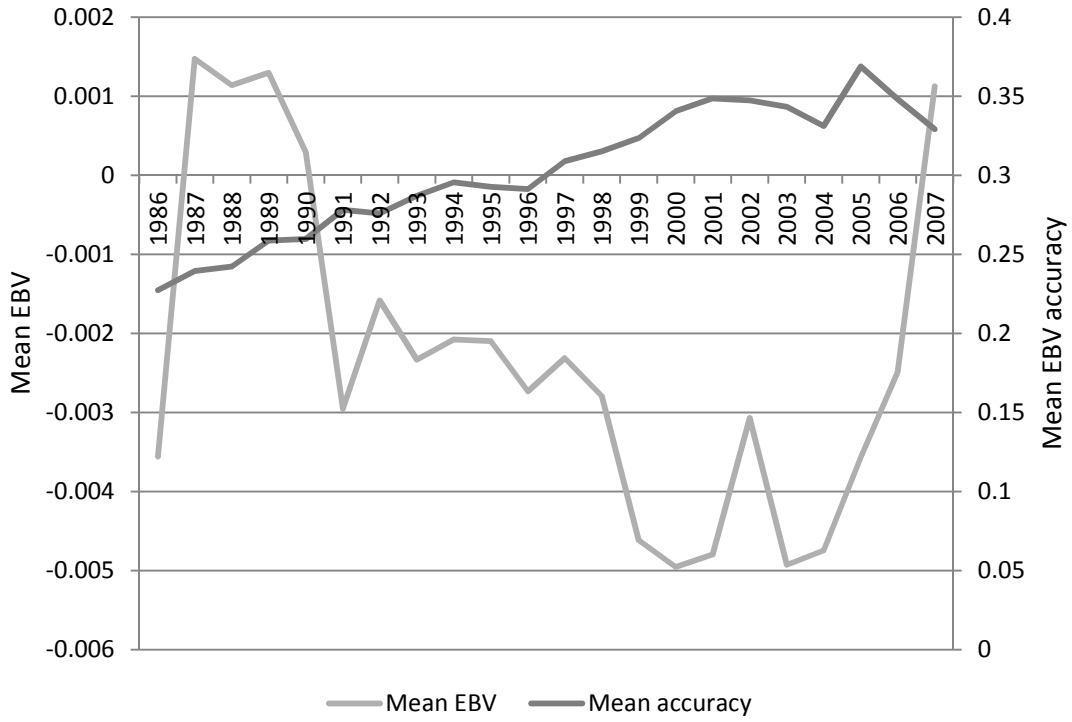

Figure 4-1. Mean EBV and mean EBV accuracy for suspensory ligament injury in Hong Kong by year of birth cohort, derived from the final linear animal model.

Both linear and logistic model EBVs show a similar trend. EBVs for suspensory ligament injury appeared to improve (become more negative) for horses born in the 1990s, but since that time, have worsened greatly, particularly since 2004. Regression of logistic EBVs on year of birth was equivalent to a decrease of -0.0023± 0.0004 grades per annum, indicating a small but significant decreasing (improving) trend overall. EBV accuracy has steadily risen over the study period.

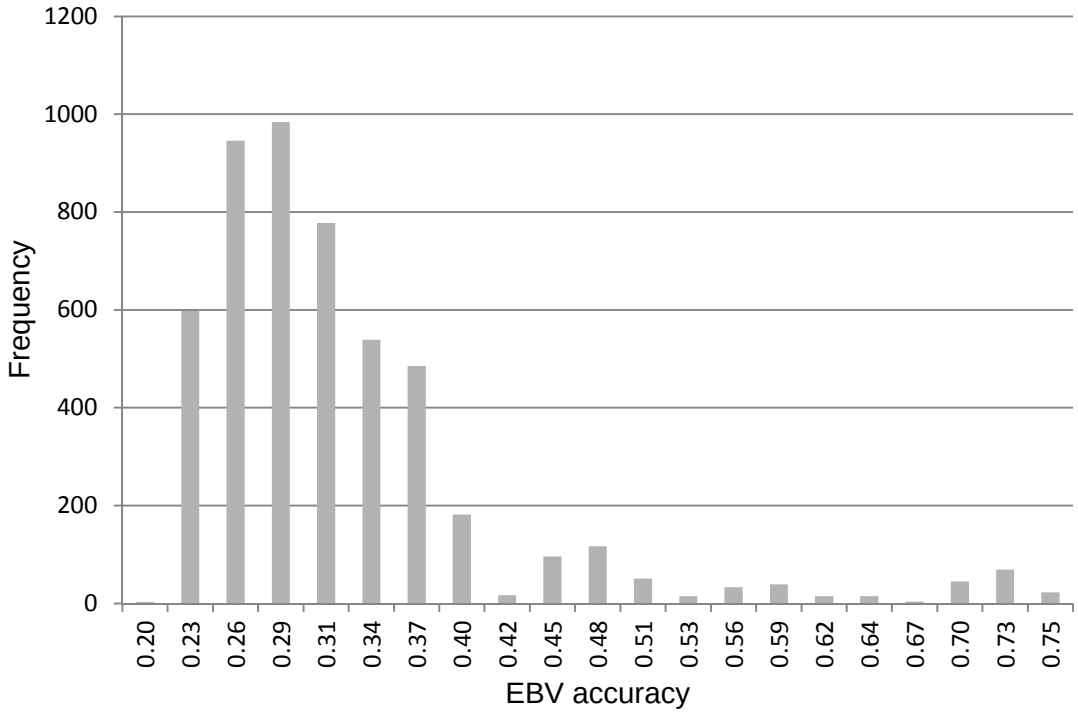

Figure 4-2. Frequency histogram of EBV accuracy for suspensory ligament injury, derived from the final linear animal model.

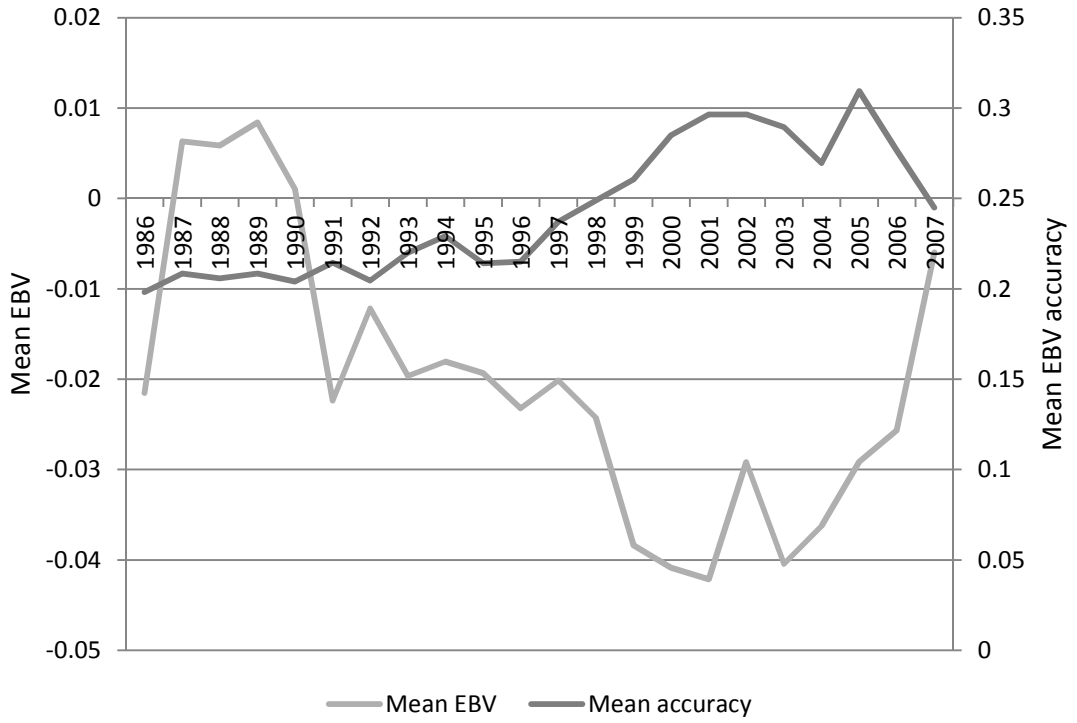

Figure 4-3. Mean EBV and mean EBV accracy for suspensory ligament injury by year of birth cohort, derived from the final logistic animal model.

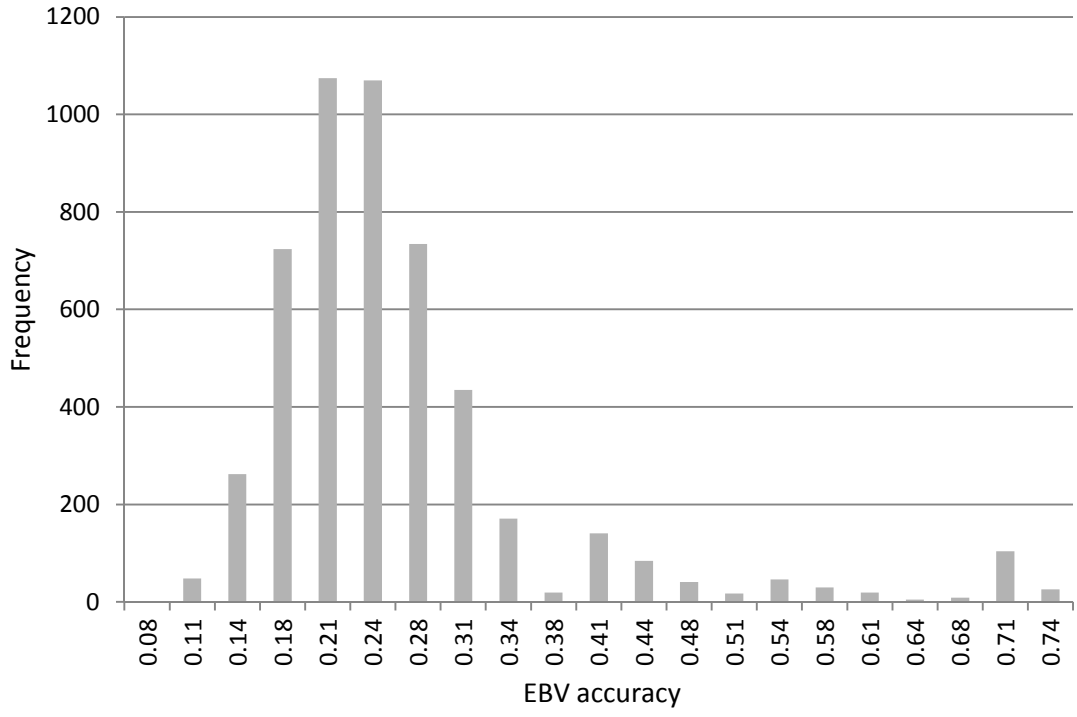

Figure 4-4. Frequency histogram of EBV accuracy for suspensory ligament injury, derived from the final logistic animal model.

Spearman rank correlation coefficient between EBVs generated using linear and logistic animal models of suspensory ligament injury was 0.992 (*p*-value <0.001), indicating good agreement between the ranking of animals based on these two model forms.

4.1.1.2 Suspensory ligament injury discussion

The prevalence of suspensory ligament injury in the current population was 10.4% (524/5062). Suspensory ligament injury accounted for 89.6% of all ligament injuries. The prevalence of suspensory ligament injury found in this study exceeds estimates from previous studies (Kasashima et al., 2004; Dyson et al., 2008). The study of Japanese Thoroughbreds by Kasashima et al. (2004) only included flat racing horses, and the UK study by Dyson et al. (2008) restricted prevalence to 2 and 3 year old horses, thus estimates of prevalence may not be directly comparable to those reported here.

Risk factors found to be associated with suspensory ligament injury in this study were the age at retirement, number of starts over the whole career and being imported from certain continents. Sex was not found to be associated with suspensory ligament injury in this study, despite the findings of previous studies in which male sex has been found to be a risk factor (Kasashima et al., 2004; Perkins et al., 2005a). Other studies have also failed to show an association between sex and risk of tendon or ligament injuries (Cogger et al., 2008; Ely et al., 2009). However, there were very few female horses in this study (0.7% of horses), thus this lack of association should be interpreted cautiously, and is more likely to represent a lack of statistical power to identify such an association, should it exist.

The positive association found here between suspensory ligament injury and age has been found in previous studies (Hill et al., 2001; Williams et al., 2001; Kasashima et al., 2004; Perkins et al., 2005a; Ely et al., 2009). These studies all found that the risk of ligament (and tendon) injuries increased with increasing age. It is unknown precisely what pathophysiological mechanisms are responsible for this association, but it is likely that it reflects either accumulated changes following exercise leading to increased susceptibility, or that normal ageing changes in ligament tissue make it more vulnerable to damage. The results of this study confirm that older racehorses should be monitored closely for signs of ligament injury due to the serious nature of the condition, and its increased likelihood in this population.

Horses that were imported from Australasia or North America had reduced odds of suspensory ligament injury in this study compared with those imported from Europe. All of these areas have their own Thoroughbred stud books, with thriving breeding industries.

Importation of racehorses to North America and Australasia is not prohibited, but it is likely that the majority of the racing population in each area originated from that area. It is likely, therefore, that horses originating from different geographical areas are, at least in part, genetically distinct from Thoroughbreds registered in different stud books. Therefore the 'continent group' variable used in the final multivariable model here may reflect different genetic groups of horses. An alternative explanation is that the environment experienced by racehorses from different areas varies, perhaps due to differences in training practices.

Table 4-1 provides evidence that suspensory ligament injury may be a heritable condition in this Thoroughbred population. Despite this, confidence intervals of estimates of heritability from logistic models spanned zero (Table 4-3). Heritability ranged from 0.05 to 0.17.

Animal model heritability estimates were smaller than those from equivalent sire models. This relationship has been found in other horse populations (Weideman et al., 2004), and in heritability analyses of other conditions in the current study. The standard errors associated with these heritability estimates were more than two times larger using sire versus animal models. For this reason it is suggested that animal models are most appropriate for modeling suspensory ligament injury as a binary trait in racehorses.

Only one statistically significant genetic correlation was found between suspensory ligament injury and another condition. Fracture was found to be positively associated (sire model) with this outcome. This finding suggests that either the genes that confer risk of these conditions are the same, or are in linkage with each other such that they are likely to be inherited together. Suspensory ligament injury was found to be significantly negatively phenotypically correlated with EIPH/epistaxis and tendon injury. No genetic correlation was found between these conditions, suggesting that the source of the negative correlation is environmental in origin. As has been previously mentioned, when multiple conditions occur in a racehorse concurrently, HKJC staff may be disinclined to enter details of both into health records, especially if the horse is to be retired. This disinclination may be because the clinical signs of one condition may be too similar to those of another condition (e.g. heat, pain and swelling associated with tendon and ligament injuries), such that diagnosis of both together would require diagnostic imaging.

Accuracies of EBVs for suspensory ligament injury were disappointing. Linear model EBV accuracies were slightly greater than those of logistic models. Estimated breeding values per year of birth cohort followed the inverse of the trend of the number of horses contributing phenotype information to each cohort (Figure 2-1). Based on these results, it is suggested that EBVs produced using the methods and data shown here would not be useful for swift selective breeding against suspensory ligament injury.

4.1.2 Tendon injury

The prevalence of tendon injury in the current dataset was 18.8% (952/5062). Records that cited tendon injury rarely noted the specific tendon injured, therefore more specific case definitions such as SDFT injury could not be modelled separately. However, given the high proportion of tendon injuries found to be injuries to the SDFT in other publications of racehorse musculoskeletal disease, it is likely that the vast majority of tendon cases reported here refer to the SDFT (Ely et al., 2004; Pinchbeck, 2004).

4.1.2.1 Multivariable analysis

Table 4-5 gives the results of likelihood ratio tests on final models of tendon injury in Hong Kong. In all models, addition of the genetic random effect (animal or sire) significantly improved the model fit.

Table 4-5. Likelihood ratio test statistics and *p*-values of final models of tendon injury. LRT statistic is determined by $2(logL_{full} - logL_{red})$ where *L* is likelihood, and *full* and *red* indicate the final model, and the 'reduced' final model, i.e. minus the genetic random variable only (sire or animal).

Model form		LRT statistic	*p*-value
Sire	Linear	16.58	<0.001
	Logistic	276.38	<0.001
Animal	Linear	32.36	<0.001
	Logistic	566.64	<0.001

Results of multivariable modeling of tendon injury as a binary trait are shown in Table 4-6. Fixed effects found to be associated with tendon injury were the intensity of racing

(number of weeks between successive races), the number of career starts, the continent group from which the horse was imported, and the length of racing career in years. The odds of tendon injury increased with increasing number of weeks between races (decreasing intensity), and longer career length. Originating from Australasia or 'other' areas compared to Europe was found to be a risk factor for tendon injury, whereas originating from North America was associated with reduced odds compared to originating from Europe. The odds of tendon injury reduced with increasing numbers of career starts.

Table 4-6. Results of multivariable sire regression models of tendon injury on linear and logistic scales.

		Linear model	**Logistic model**		
		Wald *p*-value	**Wald *p*-value**	**Odds ratio**	**95% CI**
Intensity (weeks between adjacent races)		<0.001	<0.001	1.037	1.023 – 1.052
Number of starts		<0.001	<0.001	0.950	0.938 – 0.962
Continent group	*1**	0.029	0.029	1 (REF)	
	2†			1.663	1.394 – 1.983
	3‡			0.746	0.569 – 0.979
	4§			2.282	1.419 – 3.671
Career length (yrs)		<0.001	<0.001	1.405	1.248 – 1.583

*Europe; †Austalasia; ‡North America; §other.

The heritability of tendon injury was estimated to be between 0.09 and 0.20 (Table 4-7). The largest estimate of heritability was produced using the logistic sire model. All estimates of heritability were significantly greater than zero.

Table 4-7. Heritability estimates of tendon injury in Hong Kong Thoroughbreds.

Model form		h^2	**Standard error**	**95% CI***
Sire	Linear	0.1265	0.0386	0.0508 – 0.2022
	Logistic	0.2027	0.0685	0.0684 – 0.3370
Animal	Linear	0.1423	0.0347	0.0743 – 0.2103
	Logistic	0.0879	0.0289	0.0313 – 0.1445

*Confidence interval (1.96 x standard error)

Table 4-8 shows genetic and phenotypic correlations between tendon injury and the named conditions, with the statistical significance of the genetic correlations. The only statistically significant genetic correlation was found between tendon injury and

EIPH/epistaxis. This correlation was negative in sign, and restricted to the animal model only. Tendon injury was significantly negatively phenotypically correlated with all other conditions.

Table 4-8. Genetic and phenotypic correlations (with standard errors and 95% confidence intervals) between tendon injury and the named conditions. Sire model results appear above animal model results, which are in bold type. P-values are based on LRT between the unconstrained model and the same model with the genetic correlation constrained to 0.00001. 'EIPH' also contains diagnoses of epistaxis.

	r_g*	s.e.†	95% CI‡	*p*-value	r_p§	s.e.	95% CI
EIPH	-0.2829 **-0.3574**	0.1632 **0.1394**	-0.6028 – 0.0370 **-0.6306 - -0.0842**	0.0919 **0.0145**	-0.1118 **-0.1123**	0.0140 **0.0142**	-0.1392 - -0.0844 **-0.1401 - -0.0845**
Fracture	0.0570 **-0.1479**	0.2345 **0.2127**	-0.4026 – 0.5166 **-0.5648 – 0.2690**	0.8065 **0.4976**	-0.1069 **-0.1081**	0.0140 **0.0140**	-0.1343 – -0.0795 **-0.1355 - -0.0807**
OA	0.1903 **0.0599**	0.1633 **0.1506**	-0.1298- 0.5104 **-0.2353 – 0.3551**	0.2435 **0.6892**	-0.0595 **-0.0592**	0.0142 **0.0143**	-0.0873 - -0.0317 **-0.0873 - -0.0312**
Suspensory ligament injury	-0.1302 **-0.1217**	0.2914 **0.2645**	-0.7013 – 0.4409 **-0.6401 – 0.3967**	0.6550 **0.6550**	-0.0636 **-0.0637**	0.0140 **0.0141**	-0.0910 - -0.0362 **-0.0913 - -0.0361**

*Genetic correlation, †standard error, ‡confidence interval, §phenotypic correlation.

EBVs extracted from the results of linear and logistic animal models of tendon injury are shown in Figures 4-5 to 4-8. There appeared to be a (very small) downward (improving) trend in EBVs for tendon injury as horses were born later in the study period. Regression of logistic EBVs on year of birth was equivalent to a decrease of -0.0024 ± 0.0006 grades per annum, indicating a small but significant decreasing (improving) trend. The majority of values of mean EBV were negative, indicating a reduced genetic risk relative to the population. EBV accuracies ranged from 0.37 to 0.80 (mean 0.48) from the linear model and 0.10 to 0.76 (mean 0.32) from the logistic model. Spearman rank correlation coefficient between EBVs generated using linear and logistic models was 0.975 (*p*-value <0.001).

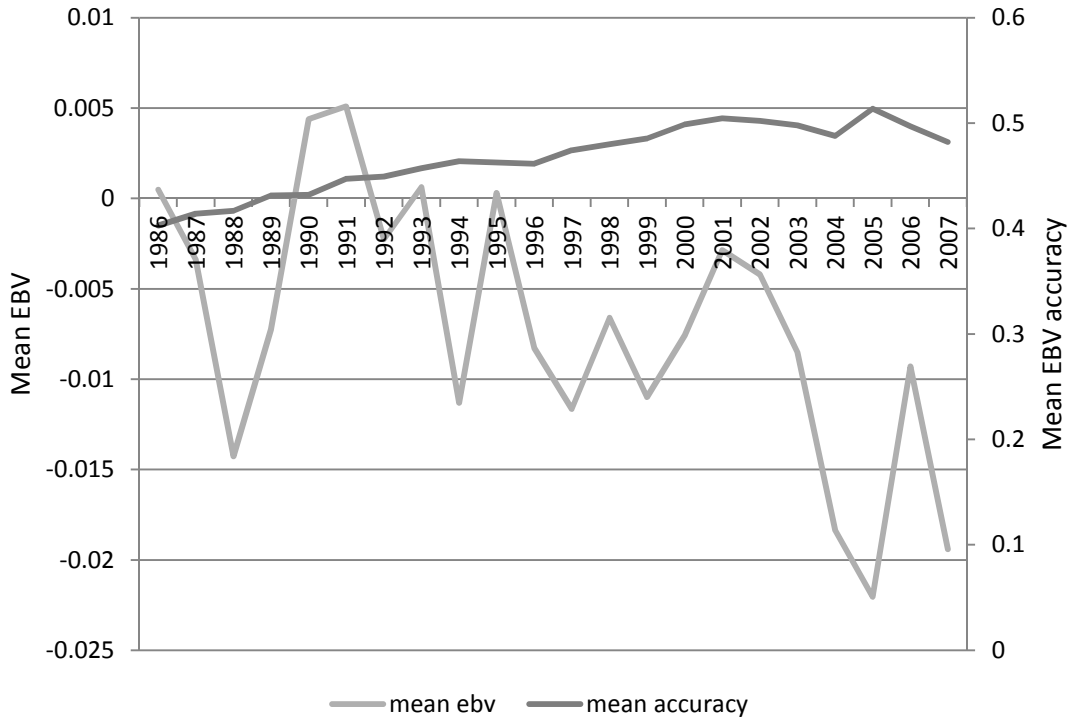

Figure 4-5. Mean EBV and mean EBV accuracies per year of birth cohort for tendon injury, derived from the final linear animal model.

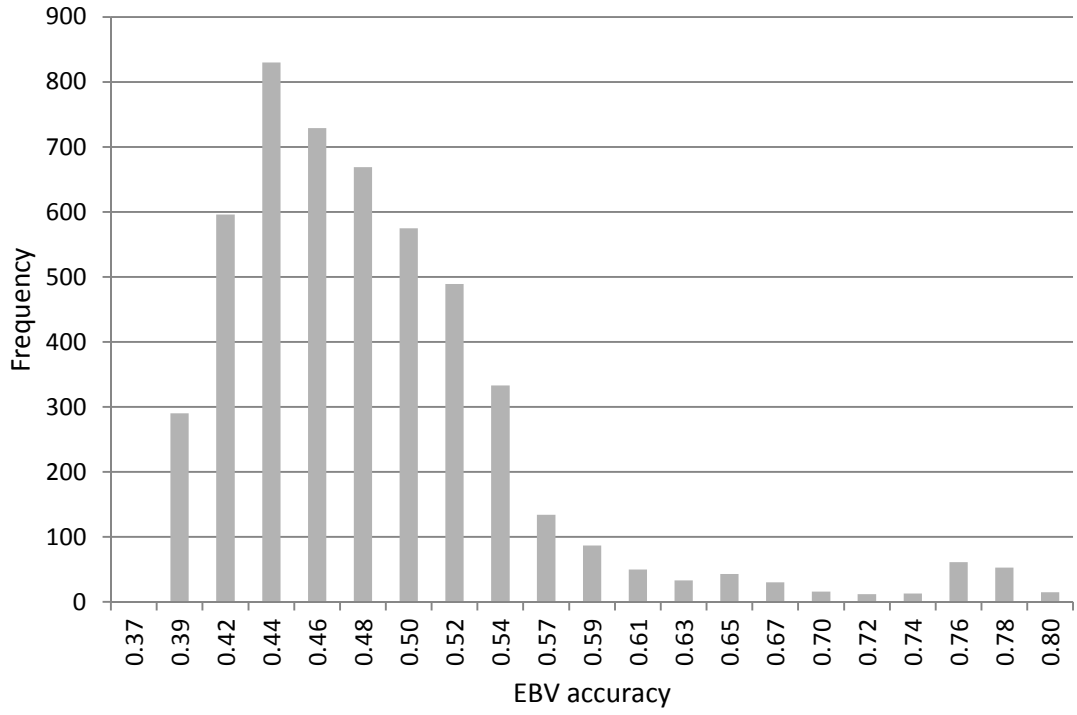

Figure 4-6. Frequency histogram of tendon injury EBV accuracies, derived from the final linear animal model.

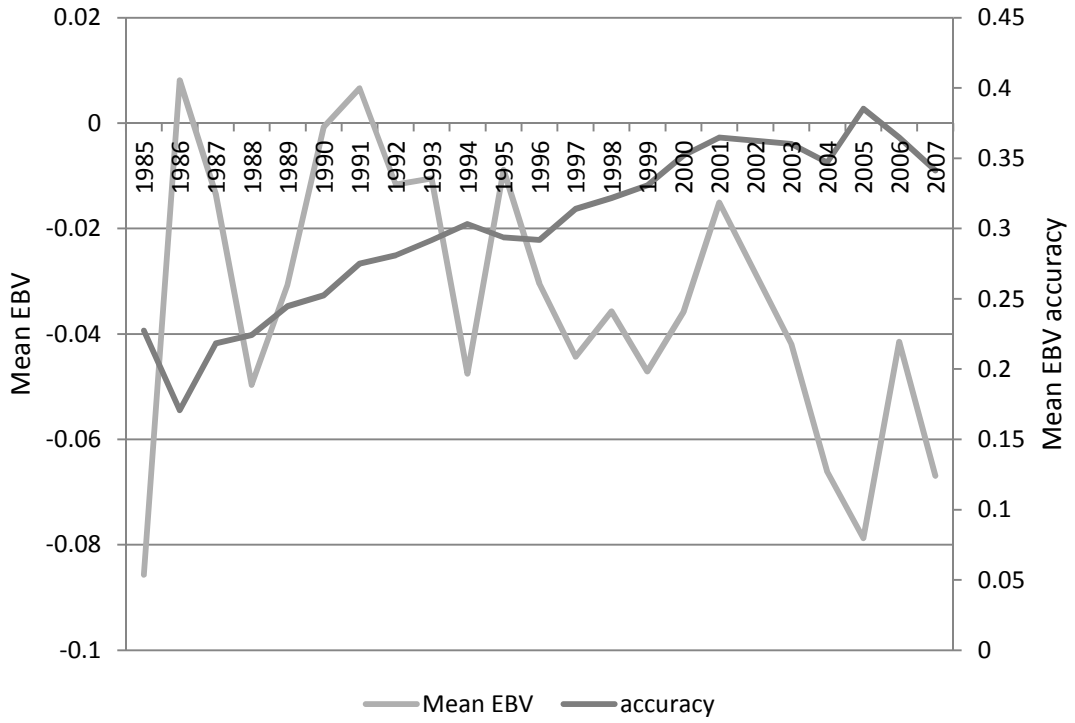

Figure 4-7. Mean EBV and mean EBV accuracies per year of birth cohort for tendon injury, derived from the final logistic animal model.

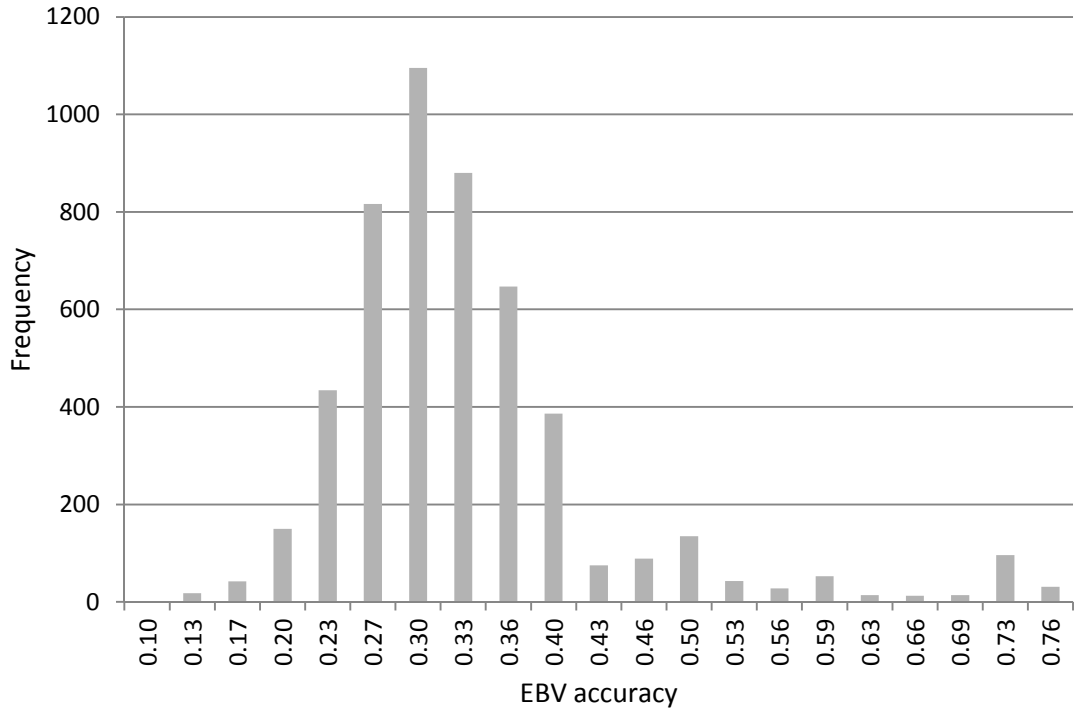

Figure 4-8. Frequency histogram of tendon injury EBV accuracies, derived from the final logistic animal model.

4.1.2.2 Tendon injury discussion

The prevalence of tendon injury in this study exceeded published figures from other populations (Lam et al., 2007b; Oki et al., 2008; Ramzan and Powell, 2010). The great majority of injuries to tendons or ligaments are injuries to the SDFT, although few studies specifically target this outcome to determine SDFT-specific risk factors and prevalence (Brommer et al., 2004; Ely et al., 2009; Neundorf et al., 2010). One study published recently examined in detail the possible environmental risk factors involved in SDFT injury in hurdle racing Thoroughbreds (Brommer et al., 2004). This study found twenty significant risk factors related to the track, race, trainer and horse. In the present study, five risk factors were found to be associated with tendon injury. As the average number of weeks between adjacent races increased (reducing intensity of racing), the odds of tendon injury increased. This finding is contrary to those from previous studies, where increasingly intense training regimes appeared to increase risk (Brommer et al., 2004; Perkins et al., 2005a; Lam et al., 2007b; Ely et al., 2009). However, the relationship between exercise history and risk is not straight forward, as running more races, over a shorter career length, appears to reduce the risk of tendon injury in the present study. These findings seem to imply that lay-ups during the racing career should be avoided. Of course, these measures of exercise history are potentially too 'crude' to reveal subtle relationships between exercise and injury as they take into account the whole racing career or a large portion of it. Assessing more recent exercise history for evidence of association with injury may be more beneficial in future studies.

Increasing age has been shown numerous times to be a risk factor for tendon injury, or injury to tendons and ligaments (Kasashima et al., 2004; Perkins et al., 2005a; Lam et al., 2007b; Oki et al., 2008; Ely et al., 2009; Neundorf et al., 2010). In the present study, age at retirement was not associated with tendon injury, but increased career length (years) was found to be a risk factor. Ageing change in tendon parenchyma may render it more susceptible to damage, but the precise mechanism behind this association is unknown (Dudhia et al., 2007; Costa et al., 2012).

As with ligament injuries, the present study failed to show an association between sex and tendon injury, as has been identified in many previous studies (Kasashima et al., 2004; Perkins et al., 2005a; Oki et al., 2008; Neundorf et al., 2010). This is likely due to the small number of female horses in the dataset.

The geographic area from which a horse originated was linked to tendon injury risk in the present study. Unlike risk of suspensory ligament injury, European horses compared to those from other areas were at a lower risk for tendon injury. Again, it is possible that different genetic groups are represented by the continent groups, but horse management differences may also play a part. Average career length, number of starts, and intensity of racing while in Hong Kong do not differ significantly between the four continent groups (data not shown), therefore differences between continents in the distribution of these variables do not appear to be responsible for this finding.

Tendon injury is a heritable condition in the population studied here. Evidence for this includes the results shown in Table 4-5, and the confidence intervals of Table 4-7. Addition of sire or animal as a random genetic effect significantly improved the fit of most of the final multivariable models for binary tendon injury, and all heritability estimates were significantly greater than zero. Estimates ranged from 0.09 to 0.20, with the sire logistic model producing the largest estimate, and the smallest confidence intervals were from the estimate based on the logistic animal model. These estimates are in close agreement with estimates of heritability of SDFT injury produced by Oki et al. (2008), based on a Gibbs sampling method in Japanese Thoroughbreds (Oki et al., 2008). As has been seen previously in this study, it is not unusual for estimates of heritability from sire models to exceed those from animal models, especially if the pedigreed population contains many full siblings, which inflate the genetic variance through inclusion of dominance effects. Unlike other conditions in this study, the heritability estimates of tendon injury were not always larger from logistic models compared with linear models. This may suggest that tendon injury may indeed be an all-or-nothing event in some cases, without a more graduated, unseen, underlying development towards complete failure. It must be remembered, however, that all types of tendon injury were included in the case definition in this part of the study, and therefore injuries to structures other than the SDFT may have altered the risk factor analyses away from true SDFT associations. Most tendon injuries in other studies were injuries to the SDFT, but a proportion involved other tendinuous structures (Ely et al., 2004; Pinchbeck, 2004). Additionally, some injuries to tendons may have been traumatic ruptures and other ‘unpredictable’ injury types, rather than exercise-related wear and tear, meaning the estimate of heritability may have been reduced. Tendon injury in the present population is of moderate heritability, which could be targeted by selective breeding strategies.

EBVs for tendon injury were moderate to poor, but showed an improving pattern in later year of birth cohorts. Linear and logistic EBV rankings were highly correlated, thus either model form could be employed to generate them, although accuracies were better for linear than for logistic model estimates. As seen in previous sections of this study, EBV accuracy improved steadily per year of birth cohort, which was most likely an effect of increasing amounts of available information in later generations.

Tendon injury, and in particular injury to the SDFT, is a serious, time-consuming outcome, with a high likelihood of re-injury due to the altered physical properties of healed tendon tissue (Brommer et al., 2004; O'Meara et al., 2010; Thorpe et al., 2010). The results presented here are in agreement with Oki et al. (2008), that reducing incidence could be attempted using selective breeding based on EBVs as selection criteria, or that molecular genetic studies are now warranted to examine the genes underlying the risk of tendon injury.

4.1.3 Osteoarthritis and joint disease

There were a total of 508 (10%, n=5062) horses diagnosed at least once with OA in the HKJC data.

4.1.3.1 Multivariable analysis

Table 4-9 shows the results of likelihood ratio tests of final multivariable models of OA in HK Thoroughbreds. Both logistic animal and sire models were significantly improved by addition of the random genetic variable, but neither of the linear models was improved.

Table 4-9. Likelihood ratio test statistics and *p*-values of final models of OA. LRT statistic is determined by $2(logL_{full} - logL_{red})$ where *L* is likelihood, and *full* and *red* indicate the final model, and the 'reduced' final model, i.e. minus the genetic random variable only (sire or animal).

Model form		LRT statistic	*p*-value
Sire	Linear	2.76	0.10
	Logistic	146.82	<0.001
Animal	Linear	-218.14	n/a
	Logistic	335.38	<0.001

Only the year of retirement was found to be associated with OA (Table 4-10). Retiring later in the study period was found to be associated with increased odds of OA.

Table 4-10. Results of multivariable sire regression models of OA on linear and logistic scales.

	Linear model	Logistic model		
	Wald *p*-value	Wald *p*-value	Odds ratio	95% CI*
Year of retirement	<0.001	<0.001	1.426	1.377 – 1.478

*Confidence interval (1.96 x standard error).

Heritability estimates of OA with their standard errors are shown in Table 4-11. Heritability ranged from 0.01 to 0.15. Confidence intervals spanned zero, but most heritability estimates exceeded their standard errors.

Table 4-11. Heritability estimates of OA in Hong Kong Thoroughbreds.

Model form		h^2	Standard error	95% CI*
Sire	Linear	0.0342	0.0250	-0.0148 – 0.0832
	Logistic	0.1461	0.0830	-0.0166 – 0.3088
Animal	Linear	0.0073	0.0154	-0.0229 – 0.0375
	Logistic	0.0539	0.0412	-0.0269 – 0.1347

*Confidence interval (1.96 x standard error)

OA was found to be positively genetically correlated with EIPH/epistaxis (animal model only) and fracture (Table 4-12). OA was also positively phenotypically correlated with fracture and suspensory ligament injury, and negatively phenotypically correlated with tendon injury.

Table 4-12. Genetic and phenotypic correlations (with standard errors and 95% confidence intervals) between OA and the named conditions. Sire model results appear above animal model results, which are in bold type. P-values are based on LRT between the unconstrained model and the same model with the genetic correlation constrained to 0.00001. 'EIPH' includes diagnoses of epistaxis.

	r_g*	s.e.†	95% CI‡	*p*-value	r_p§	s.e.	95% CI
EIPH	0.2344 **0.2732**	0.1347 **0.1333**	-0.0296 – 0.4984 **0.0119 – 0.5345**	0.0773 **0.0368**	-0.0081 **-0.0106**	0.0143 **0.0144**	-0.0361 – 0.0199 **-0.0388 – 0.0176**
Fracture	0.8499 **0.8930**	0.1347 **0.1422**	0.5859 – 1.1139 **0.6143 – 1.1717**	<0.001 **<0.001**	0.1682 **0.1662**	0.0138 **0.0138**	0.1412 - 0.1952 **0.1392 – 0.1932**
Suspensory ligament injury	0.1162 **0.1326**	0.2426 **0.2389**	-0.3593 – 0.5917 **-0.3356 – 0.6008**	0.6390 **0.5720**	0.0333 **0.0339**	0.0141 **0.0141**	0.0057 – 0.0609 **0.0063 – 0.0615**
Tendon injury	0.1903 **0.0599**	0.1633 **0.1506**	-0.1298 – 0.5104 **-0.2353 – 0.3551**	0.2435 **0.6892**	-0.0595 **-0.0592**	0.0142 **0.0143**	-0.0873 - -0.0317 **-0.0872 - -0.0312**

*Genetic correlation, †standard error, ‡confidence interval, §phenotypic correlation.

Mean EBV per year of birth cohort, and mean EBV accuracy are shown in Figures 4-9 and 4-11. Figures 4-10 and 4-12 show frequency histograms of EBV accuracies for OA over the study period. These figures show a very small increasing (worsening) trend in EBVs over the period in question, with the majority of mean EBV values lying above zero. Regression of logistic EBVs on year of birth was equivalent to an increase of 0.0007 ± 0.0001 grades per annum, indicating a small but significant increasing (worsening) trend. EBV accuracy appeared to increase steadily over the study period. Spearman rank correlation coefficient between linear and logistic model EBVs was 0.975 (*p*-value <0.001). Linear model EBV accuracies ranged from 0.03 to 0.70 (mean 0.08), and logistic model EBV accuracies ranged from 0.02 to 0.72 (mean 0.14).

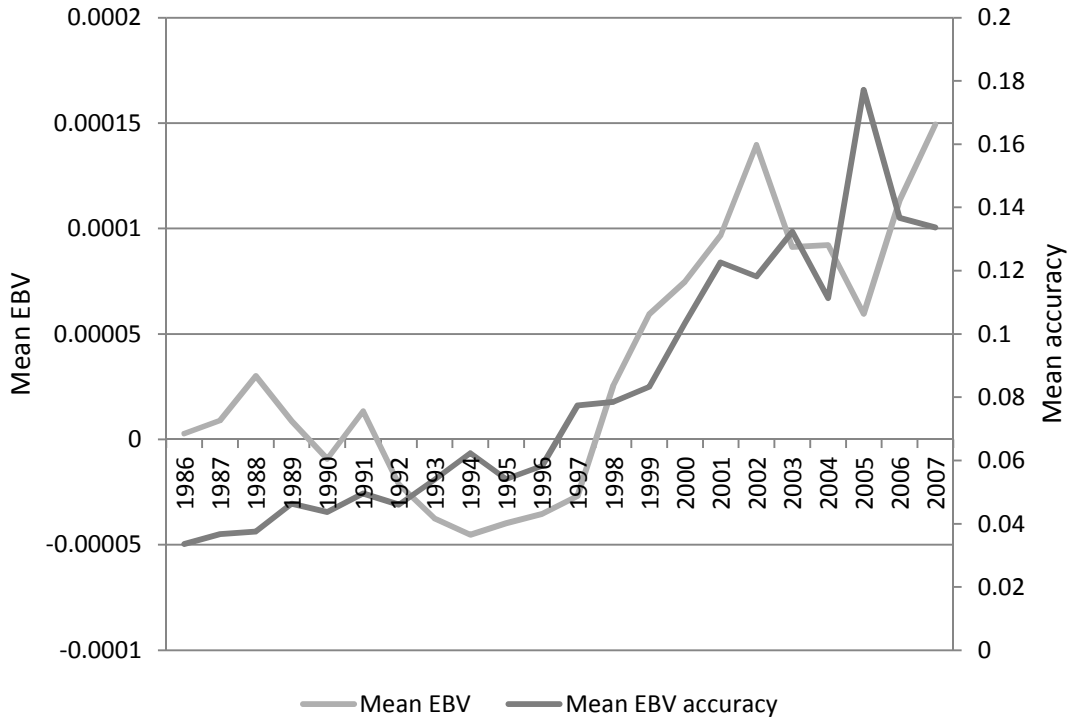

Figure 4-9. Mean EBV and mean EBV accuracies per year of birth cohort for OA, derived from the final linear animal model.

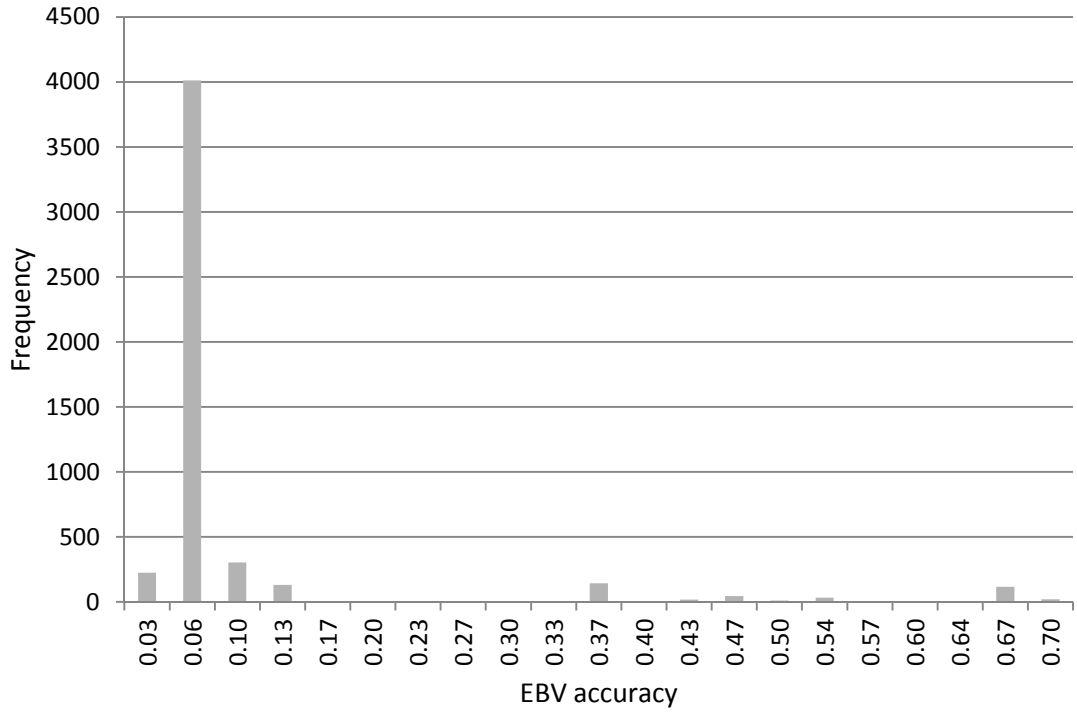

Figure 4-10. Frequency histogram of OA EBV accuracies, derived from the final linear animal model.

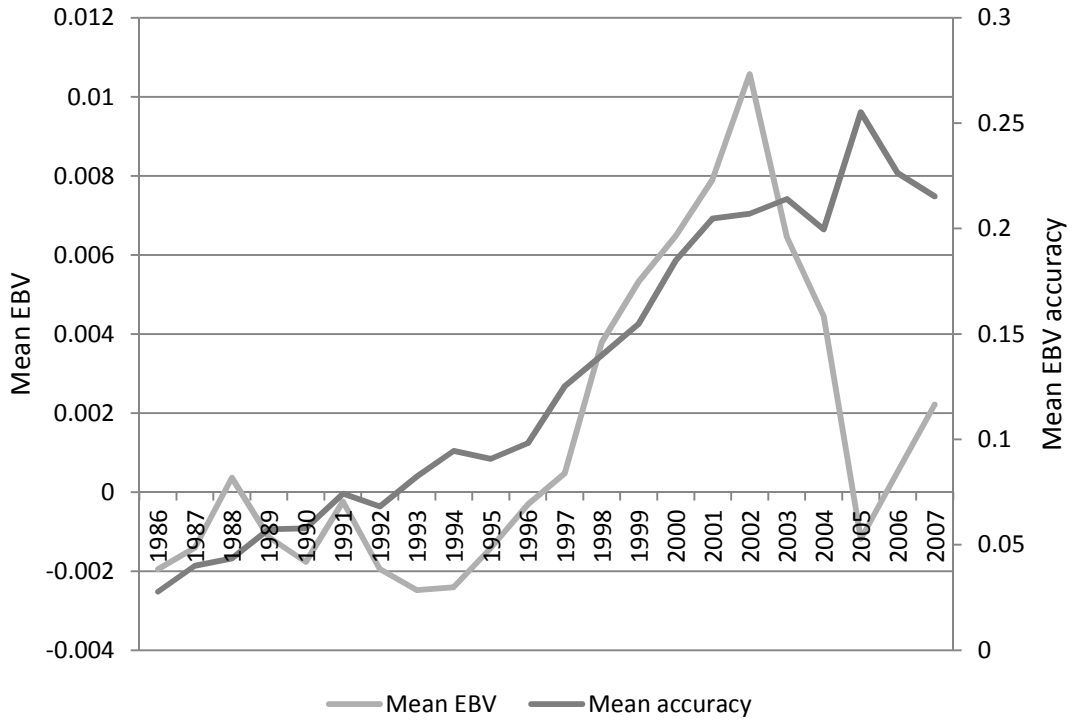

Figure 4-11. Mean EBV and mean EBV accuracies per year of birth cohort for OA, derived from the final logistic animal model.

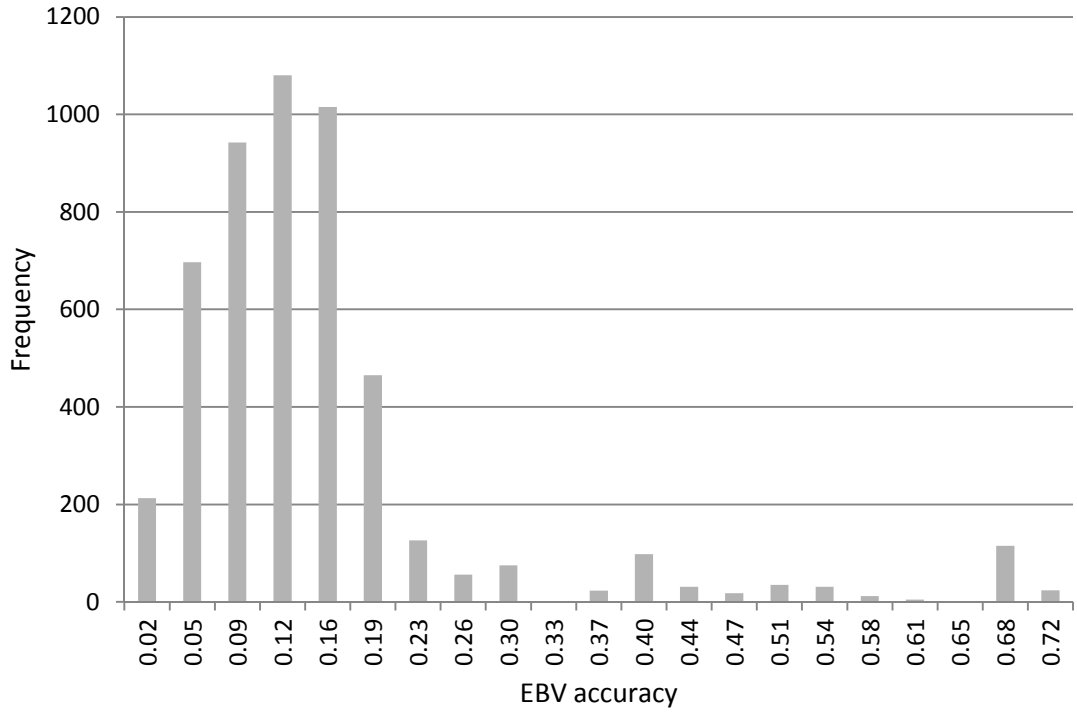

Figure 4-12. Frequency histogram of OA EBV accuracies, derived from the final logistic animal model.

4.1.3.2 Joint disease multivariable analysis

The OA and DJD status of each horse were combined into a single binary variable called 'joint disease', which was coded '1' if a horse was ever diagnosed with OA or DJD or both (n=749, 15%), and 0 if that horse had never been diagnosed with either condition. This disease category was modelled in the same way as all other conditions, and the results of this modelling are shown below. Table 4-13 gives the likelihood ratio test results for joint disease final models, and Table 4-14 gives the results of fixed effects found to be significantly associated with joint disease. Similarly to OA, only logistic models of joint disease were improved by the addition of a genetic random variable.

Table 4-13. Likelihood ratio test statistics and *p*-values of final models for joint disease. LRT statistic is determined by $2(logL_{full} - logL_{red})$ where *L* is likelihood, and *full* and *red* indicate the final model, and the 'reduced' final model, i.e. minus the genetic random variable only (sire or animal).

Model form		LRT statistic	*p*-value
Sire	Linear	2.26	0.122
	Logistic	173.44	<0.001
Animal	Linear	0.88	0.348
	Logistic	248.20	<0.001

Two fixed effects were found to be associated with joint disease; year of retirement and age at retirement (years). The odds of joint disease increased with later year of retirement, and older age at retirement.

Table 4-14. Results of mutivariable sire regression models of joint disease on linear and logistic scales.

	Linear model	Logistic model		
	Wald *p*-value	Wald *p*-value	Odds ratio	95% CI
Year of retirement	<0.001	<0.001	1.134	1.110 – 1.158
Age at retirement (yrs)	<0.001	<0.001	1.275	1.218 – 1.335

Heritability estimates for joint disease are shown in Table 4-15. All confidence intervals for heritability estimates spanned zero. The heritability of joint disease ranged from 0.02 to 0.11. The highest estimate of heritability was produced by the logistic sire model.

Table 4-15. Heritability estimates of joint disease in Hong Kong Thoroughbreds.

Model form		h^2	Standard error	95% CI*
Sire	Linear	0.0398	0.0285	-0.0161 – 0.0957
	Logistic	0.1080	0.0688	-0.0269 – 0.2428
Animal	Linear	0.0150	0.0184	-0.0211 – 0.0511
	Logistic	0.0312	0.0318	-0.0311 – 0.0935

*Confidence interval (1.96 x standard error)

4.1.3.3 Osteoarthritis and joint disease discussion

At 10% of horses affected, the prevalence of OA in this population is somewhat less that has been previously found in other Thoroughbred populations. However, previous studies tended to use radiographic diagnosis rather than clinical signs, therefore some subclinical OA cases will have been missed in this study (Neundorf et al., 2010). Also, some horses suffering from OA may have been retired previously for other reasons, and in some cases the OA detected may not have been severe enough to warrant an OVE or lead to retirement. It is not clear whether terminology relating to OA, and to degenerative joint disease (DJD), were used interchangeably by HKJC vets in diagnosis of joint disease. Many equine clinicians use the terms OA and DJD as synonyms. However, DJD is occasionally reserved for cases of primary OA, where no inciting cause can be found (McIlwraith and Vachon, 1988). For this reason, OA and DJD were modelled separately in this population initially. The prevalence of DJD was 5%, and due to this low case number, model convergence was not achieved (data not shown). Results of a bivariate animal model devoid of fixed effects between OA and DJD revealed a significant negative genetic correlation between these diagnoses (r_g -0.74, s.e. 0.13, *p*-value <0.001), but no significant phenotypic correlation. These results suggest that OA and DJD are genetically different entities in this population, if they have been classified correctly. Nevertheless, incase DJD and OA terms were used to describe the same trait, both were combined into a disease category called 'joint disease'. For a horse to be positive for joint disease in the HKJC dataset, it must have been diagnosed at least once with either OA, or DJD, or both. There were 749 (15%) horses with joint disease. Heritability estimates of joint disease were larger than estimates of OA heritability, but the differences were not significant. Both OA and joint disease were associated with the year of retirement. These similarities suggest a commonality between joint disease and OA. Genetic correlations and EBVs were not estimated for joint disease due to non-significant heritability estimates.

This study provides some evidence that OA (DJD) is a heritable condition in this population, although this evidence is somewhat weak. Osteoarthritis is a term meaning inflammation of the bone and joint tissues, which can arise for many different reasons. For example, inflammation can be caused by trauma, sepsis, auto-immunity, or as a result of developmental orthopaedic conditions such as OCD. The equine metacarpophalangeal joint is known to be at greatest risk of OA, probably because of the stresses this joint is subjected to during exercise (Neundorf et al., 2010). In most cases, the inciting cause of the OA (DJD) diagnosed in the HKJC data was not known or not specified, and the joint or limb involved was not always stated, therefore the case definition contained all sites and causes of OA (DJD). In humans, OA is known to have a strong genetic component, which appears to vary upon ethnicity and the joint(s) involved. However, it is possible that most OA (DJD) seen in the current equine population was secondary to other more prevalent conditions such as OCD, rather than being primary. No records from the HKJC contained diagnoses of OCD. Horses were imported into Hong Kong at an average age of 2.8 years (range 1.7 to 7.0, mean 2.8, median 2.8, mode 2.7 years). OCD is known as a developmental disorder, and therefore may be reserved for diagnoses of arthropathy in young horses, therefore the joint pathology diagnosed by the HKJC team would have been more likely to be classified as OA (DJD). If this is the case, then heritability estimates of joint disease or OA shown here would be artificially low. Heritability of primary OA in this population would be being masked by secondary cases, and the heritability of OCD would be only partially captured through cases of secondary OA. More specific case definitions, probably with the use of visual diagnostic aids such as radiography, would be required in future studies to ascertain the nature of the heritable component of joint disease shown here.

4.1.4 Hong Kong musculoskeletal condition discussion

Musculoskeletal disorders are prevalent in the Hong Kong Thoroughbred population. This study provides some evidence that a number of these conditions are, in part, heritable, and are in many cases positively genetically correlated with each other, suggesting that breeding strategies for reducing incidence may affect a number of conditions concurrently.

Past studies have often tended to combine injuries to tendons and ligaments into a single case definition, and study the risk factors for these injuries combined (TLI). This may be the case because of the proximity of the digital flexor tendons to the suspensory ligament on the palmar/plantar aspect of the distal limbs, making ultrasonographic interpretation of the health of these structures easy to assess during the same examination. Similar clinical signs are involved when damage to these ligaments or tendons occurs. However, tendons and ligaments occupy different anatomical locations and structural properties, and are important to locomotion and joint stability in different ways. The results presented here show that the risk factors associated with injuries to tendons, and to the suspensory ligament are different, and a positive genetic correlation does not exist between them. These findings suggest that combination of tendon and ligament injuries into a single dependent variable in epidemiological and genetic studies of risk factors is inappropriate.

EBV accuracies for each of the musculoskeletal conditions studied here were small or moderate in size. Selection using EBVs with these accuracies would not deliver swift reductions in genetic risk. To improve selection based on this method, more specific and accurate diagnoses would be required, with larger numbers of case and control horses than were available in the current study.

4.2 Musculoskeletal disorders in the BHA dataset

4.2.1 BHA model results using all horses

Two thousand and seventy horses (2.6%, 2070/78151) were diagnosed with at least one occurrence of injury to the SDFT in the BHA data. Of these horses, 81% (1680) were male. There were 49660 horses that competed in any flat racing, and 40135 that competed in any jump racing. Of the horses that competed in flat racing, 688 (1.4%) were diagnosed with an SDFT injury, compared to 1928 (4.8%) jump racing horses.

4.2.1.1 PH dataset SDFT injury multivariable analysis

The results of multivariable models of SDFT injury in the BHA 'per horse' (PH) data are shown in Tables 4-16, 4-17 and 4-18. Table 4-16 shows the results of likelihood ratio tests

of SDFT injury final models. Addition of the genetic random variable (sire or animal) significantly improved the fit of all models.

Table 4-16. Likelihood ratio test statistics and *p*-values of final models for SDFT injury. LRT statistic is determined by $2(logL_{full} - logL_{red})$ where L is likelihood, and *full* and *red* indicate the final model, and the 'reduced' final model, i.e. minus the genetic random variable only (sire or animal).

Model form		**LRT statistic**	***p*-value**
Sire	Linear	139.28	<0.001
	Logistic	11617.62	<0.001
Animal	Linear	157.90	<0.001
	Logistic	11861.68	<0.001

Four fixed effects were found to be significantly associated with SDFT injury in the BHA data (Table 4-17). Competing in any jump racing was strongly associated with increased odds of SDFT injury. Increasing number of career starts was also associated with increased odds, whereas being female compared to male, and accruing more winnings over the career were found to be associated with reduced odds of SDFT injury.

Table 4-17. Results of multivariable sire regression models of SDFT injury on linear and logistic scales.

		Linear model	**Logistic model**		
		Wald *p*-value	**Wald *p*-value**	**Odds ratio**	**95% CI**
Racing profile	*1**	<0.001	<0.001	1 (REF)	
	2†			8.523	6.944 – 10.459
Sex	Male	<0.001	<0.001	1 (REF)	
	Female			0.306	0.204 – 0.459
Number of starts		<0.001	<0.001	1.011	1.008 – 1.014
Winnings (thousands)		0.019	<0.001	0.995	0.993 – 0.998
Interaction 1‡		<0.001	<0.001	2.334	1.527 – 3.567

*Horses that competed exclusively in flat races; † Horses that competed in any jump races; ‡First order interaction between racing profile and sex.

4.2.1.2 PH dataset SDFT injury heritability and genetic correlation estimation

Heritability estimates of SDFT injury ranged from 0.04 to 0.24 (Table 4-18), and all estimates were significantly greater than zero. The largest heritability estimate was generated from the logistic sire model.

Table 4-18. Heritability estimates of SDFT injury.

Model form		h^2	Standard error	95% CI*
Sire	Linear	0.0385	0.0049	0.0289 – 0.0481
	Logistic	0.2373	0.0390	0.1608 – 0.3137
Animal	Linear	0.0450	0.0052	0.1395 – 0.2207
	Logistic	0.1372	0.0222	0.0937 – 0.1807

*Confidence interval (1.96 x standard error)

Table 4-19 contains results of bivariate analyses within the BHA PH dataset. SDFT injury was found to be significantly positively genetically correlated with both distal limb fracture, and epistaxis. SDFT injury was also negatively phenotypically correlated with distal limb fracture, but not with epistaxis.

Table 4-19. Genetic and phenotypic correlations (with standard errors and 95% confidence intervals) between SDFT injury, distal limb fracture and epistaxis. Sire model results appear above animal model results, which are in bold type. P-values are based on LRT between the unconstrained model and the same model with the genetic correlation constrained to 0.00001.

	r_g*	s.e.†	95% CI‡	*p*- value	r_p§	s.e.	95% CI‡
Distal limb fracture	0.4661 **0.4268**	0.0996 **0.0961**	0.2709 – 0.6613 **0.2384 – 0.6152**	0.0003 **0.0001**	-0.0140 **-0.0141**	0.0036 **0.0036**	-0.0211 – -0.0069 **-0.0212 - -0.0070**
Epistaxis	0.2942 **0.3021**	0.0748 **0.0697**	0.1476 – 0.4408 **0.1655 – 0.4387**	0.0013 **0.0001**	0.0067 **0.0066**	0.0036 **0.0036**	-0.0004 – 0.0138 **-0.0005 – 0.0137**

*Genetic correlation, †standard error, ‡confidence interval, §phenotypic correlation.

4.2.1.3 BHA PH dataset SDFT injury estimated breeding values

EBVs were extracted from the results of final SDFT injury animal models on the linear and logistic scales. Mean EBV and mean EBV accuracy for each year of birth cohort are shown in Figures 4-13 and 4-15. Histograms of EBV accuracies are shown in Figures 4-14 and 4-16. In both linear and logistic models, the mean EBV per year of birth cohort appeared to improve (become more negative) over the study period. Regression of logistic EBVs on year of birth was equivalent to a decrease of -0.0127 ± 0.0002 grades per annum,

indicating a small but significant decreasing (improving) trend. Mean accuracies increased somewhat over this period, but logistic model EBV accuracies showed less of an improvement than those from linear models. EBV accuracies ranged from 0.21 to 0.90 (mean 0.41) and 0.01 to 0.62 (mean 0.26) in linear and logistic models, respectively. Spearman rank correlation coefficient between linear and logistic EBVs was 0.956 (*p*-value <0.001).

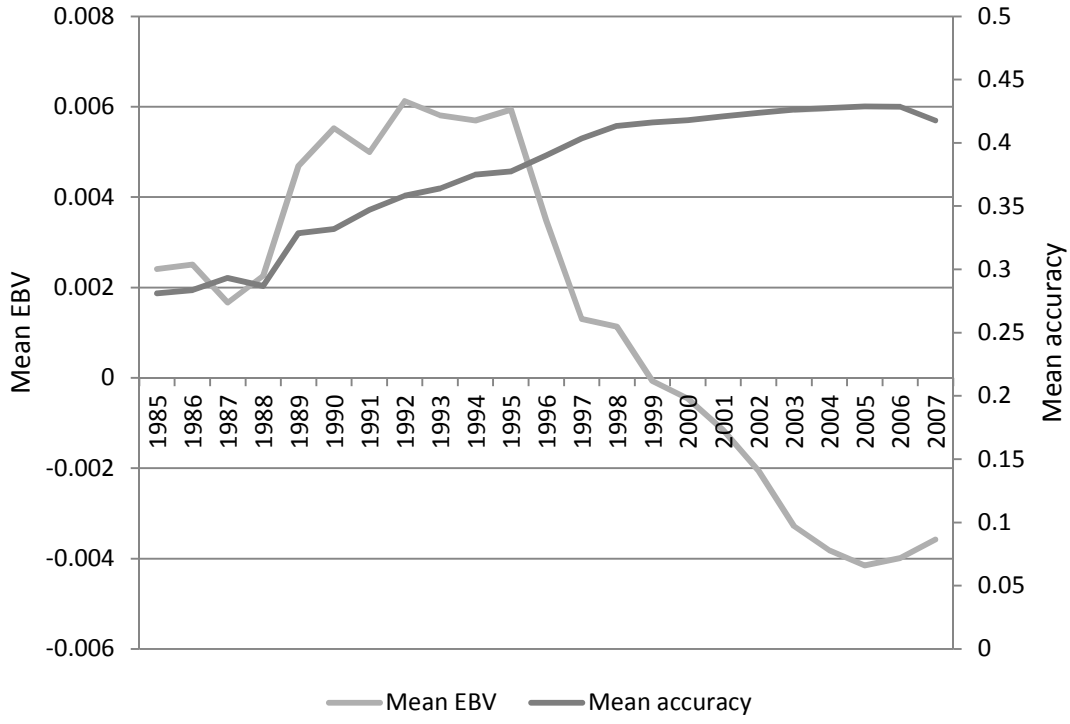

Figure 4-13. Mean EBV and mean EBV accuracies per year of birth cohort for SDFT, derived from the final linear animal model.

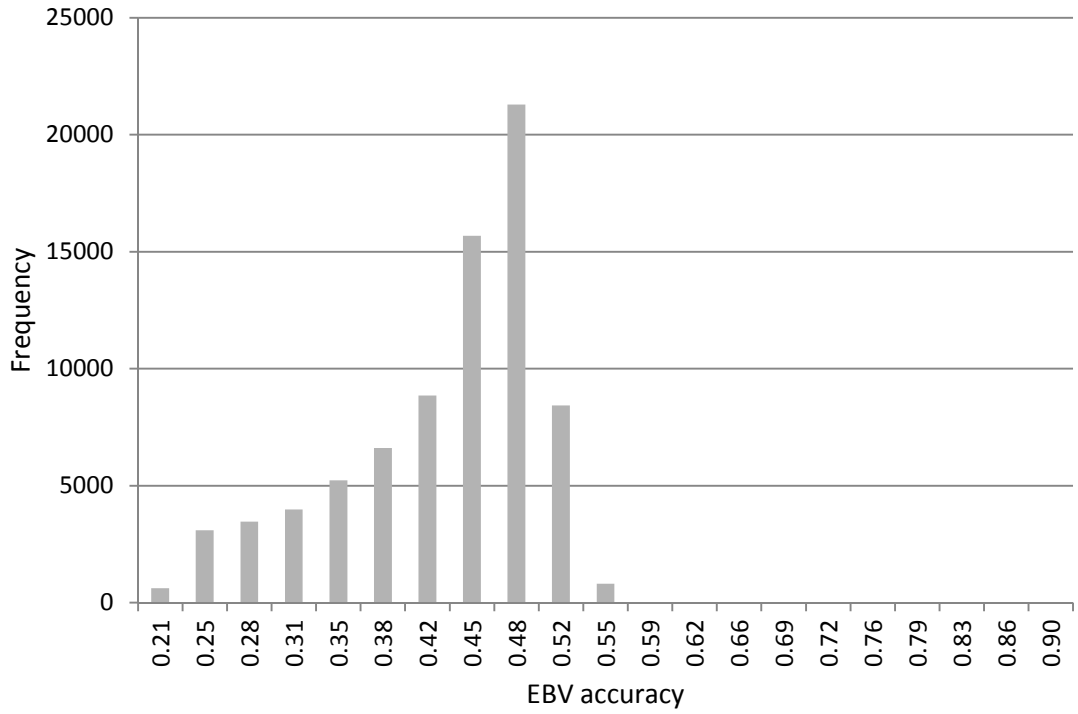

Figure 4-14. Frequency histogram of SDFT EBV accuracies, derived from the final linear animal model.

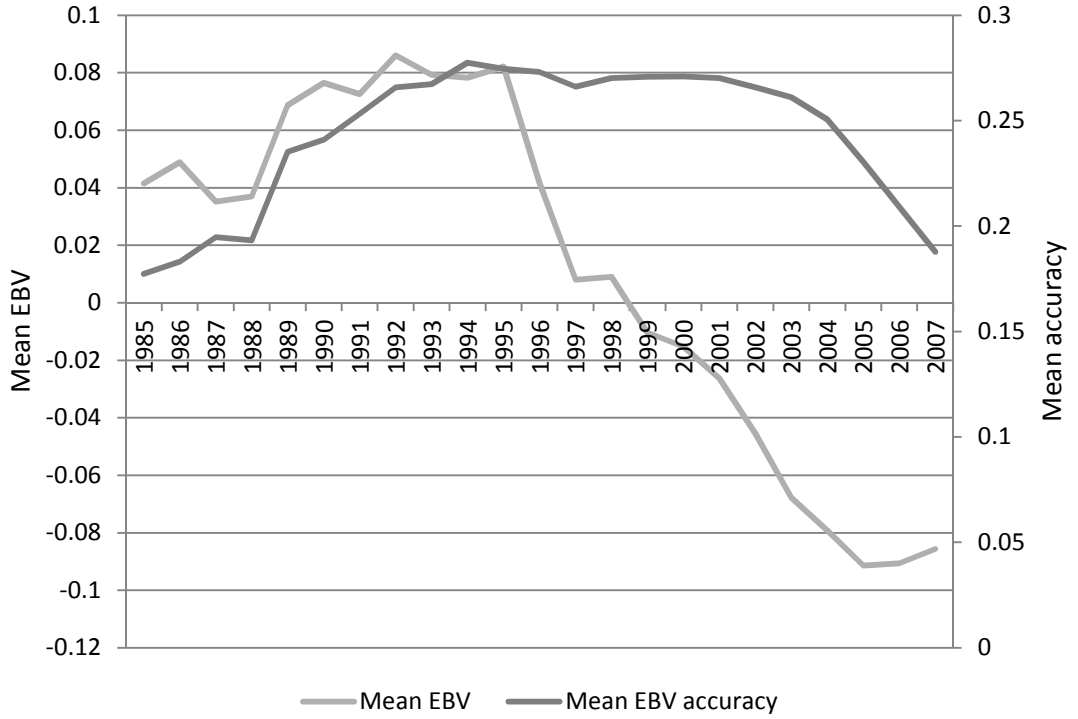

Figure 4-15. Mean EBV and mean EBV accuracies per year of birth cohort for SDFT, derived from the final logistic animal model.

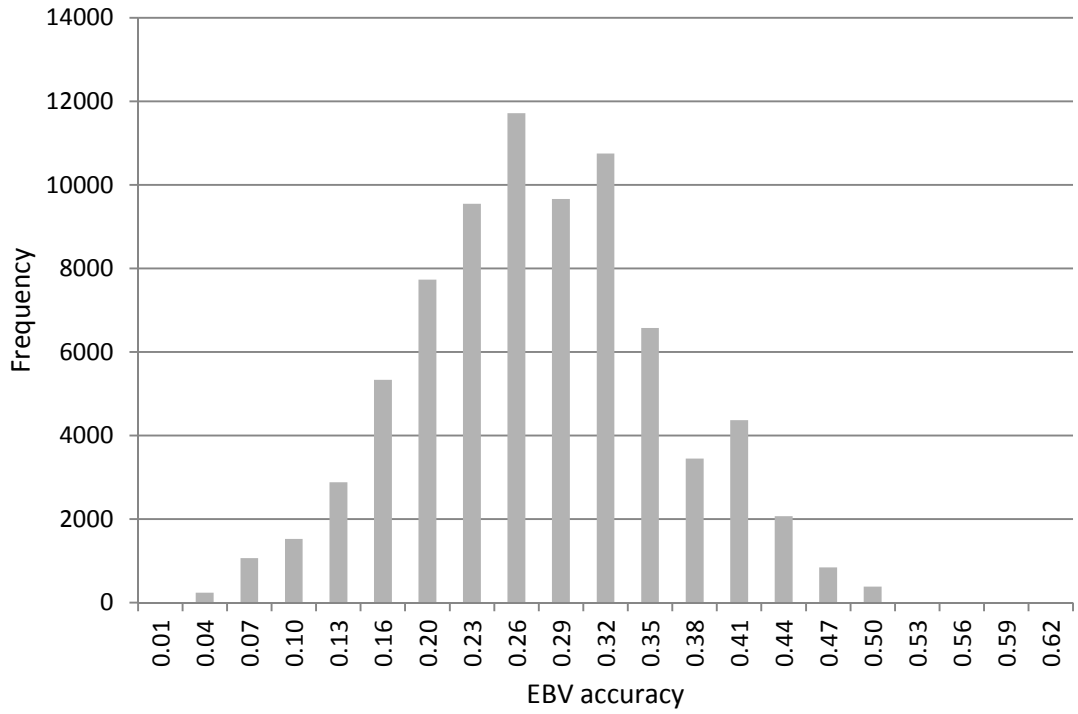

Figure 4-16. Frequency histogram of SDFT EBV accuracies, derived from the final logistic animal model.

4.2.1.4 BHA REP dataset heritability and repeatability of SDFT injury

Table 4-20 contains results of linear repeatability models of SDFT injury in the BHA repeated records (REP) dataset. Heritability accounted for 23% and 24% of repeatability in sire and animal models, respectively, with permanent environmental effects accounting for 94% and 76% of repeatability.

Table 4-20. Heritability, repeatabilty and permanent environmental variance of SDFT injury.

	h^2* (s.e.)	σ^2_{pe} † (s.e.)	r‡ (s.e.)
Sire	0.0145 (0.0013)	0.0589 (0.0008)	0.0625 (0.0008)
Animal	0.0150 (0.0012)	0.0476 (0.0012)	0.0626 (0.0008)

*Heritability;† permanent environmental variance;‡ repeatability.

The REP dataset was divided into two separate datasets, one containing all records from horses that competed in at least one flat race, and the other containing all records from horses that competed in at least one jump race. The results of repeatability models of SDFT injury within these two datasets are shown in Table 4-21. Permanent environmental effects accounted for 74% and 94% of variation in SDFT injury using sire models in flat and jump racing populations, respectively, and 46% and 74% using animal models.

Table 4-21. Heritability, repeatabilty and permanent environmental variance of SDFT injury, in both flat (upper) and jump (lower) racing populations.

	h^2* (s.e.)	σ^2_{pe} † (s.e.)	r‡ (s.e.)
Sire	0.1071 (0.0066) 0.0114 (0.0014)	0.0780 (0.0010) 0.0449 (0.0009)	0.1047 (0.0018) 0.0477 (0.0009)
Animal	0.0523 (0.0024) 0.0125 (0.0014)	0.0440 (0.0020) 0.0353 (0.0015)	0.0963 (0.0012) 0.0478 (0.0009)

*Heritability;† permanent environmental variance;‡ repeatability.

4.2.2 Flat racing population

4.2.2.1 Multivariable analysis

SDFT injury was modelled within the flat and jump racing populations in the BHA PH dataset separately. Table 4-21 shows results of likelihood ratio tests of the final SDFT models in the flat racing population. Addition of a genetic random variable improved the fit of all models except the logistic sire model.

Table 4-22. Likelihood ratio test statistics and *p*-values of final models for SDFT injury in flat racing horses only. LRT statistic is determined by $2(logL_{full} - logL_{red})$ where L is likelihood, and *full* and *red* indicate the final model, and the 'reduced' final model, i.e. minus the genetic random variable only (sire or animal).

Model form		LRT statistic	*p*-value
Sire	Linear	96.16	<0.001
	Logistic	-9115.16	n/a
Animal	Linear	105.66	<0.001
	Logistic	25537.60	<0.001

Only two fixed effects were found to be associated with SDFT injury in flat racing horses; career length (years) and sex (Table 4-23). Longer careers were associated with increased odds of SDFT injury, and being female compared to male was associated with reduced odds.

Table 4-23. Results of multivariable sire regression models of SDFT injury on linear and logistic scales.

		Linear model	Logistic model		
		Wald *p*-value	**Wald *p*-value**	**Odds ratio**	**95% CI**
Career length (yrs)		<0.001	<0.001	1.406	1.356 – 1.457
Sex	Male	<0.001	<0.001	1 (REF)	
	Female			0.480	0.388 – 0.594

Estimates of the heritability of SDFT injury in the flat racing population ranged from 0.05 to 0.52, and all estimates were significantly greater than zero (Table 4-24). Logistic models produced larger heritability estimates compared to linear models. The largest heritability estimate was generated by the logistic sire model.

Table 4-24. Heritability estimates of SDFT injury in the flat racing BHA population.

Model form		h^2	**Standard error**	**95% CI***
Sire	Linear	0.0491	0.0072	0.0350 – 0.0632
	Logistic	0.5161	0.0932	0.3335 – 0.6987
Animal	Linear	0.0497	0.0067	0.0366 – 0.0628
	Logistic	0.2460	0.0423	0.1631 – 0.3288

*Confidence interval (1.96 x standard error)

4.2.3 Jump racing population

4.2.3.1 Multivariable analysis

Likelihood ratio test results of final models of SDFT injury in the jump racing population are shown in Table 4-25. All models were significantly improved by the addition of a genetic random variable.

Table 4-25. Likelihood ratio test statistics and *p*-values of final models for SDFT injury in jump racing horses only. LRT statistic is determined by $2(logL_{full} - logL_{red})$ where L is likelihood, and *full* and *red* indicate the final model, and the 'reduced' final model, i.e. minus the genetic random variable only (sire or animal).

Model form		LRT statistic	*p*-value
Sire	Linear	51.48	<0.001
	Logistic	2810.98	<0.001
Animal	Linear	55.94	<0.001
	Logistic	10057.60	<0.001

Three fixed effects were found to be associated with SDFT injury in the jump racing population; career length (years), sex and the number of starts (Table 4-26). Both female sex and increased number of starts were associated with reduced odds of SDFT injury. Longer career length was associated with increased odds of SDFT injury.

Table 4-26. Results of multivariable sire regression models of SDFT injury on linear and logistic scales.

		Linear model	Logistic model		
		Wald *p*-value	Wald *p*-value	Odds ratio	95% CI
Career length (yrs)		<0.001	<0.001	1.272	1.229 – 1.316
Sex	Male	<0.001	<0.001	1 (REF)	
	Female			0.815	0.721 – 0.922
Number of starts		<0.001	<0.001	0.982	0.977 – 0.988

Table 4-27 shows the heritability estimates of SDFT injury in the jump racing population of the BHA data. Heritability estimates ranged from 0.03 to 0.17, all of which were significantly greater than zero. The logistic sire model produced the largest heritability estimate.

Table 4-27. Heritability estimates of SDFT injury from the jump racing BHA population.

Model form		h^2	Standard error	95% CI*
Sire	Linear	0.0289	0.0059	0.0173 – 0.0405
	Logistic	0.1739	0.0363	0.1027 – 0.2451
Animal	Linear	0.0309	0.0059	0.0193 – 0.0425
	Logistic	0.1085	0.0225	0.0644 – 0.1527

*Confidence interval (1.96 x standard error)

4.2.4 BHA SDFT injury discussion

The prevalence of SDFT injury at the racecourse reported here is lower than reports of SDFT injury in literature (Oki et al., 2008; Ramzan and Palmer, 2011). However, these studies included SDFT injuries diagnosed during training activities, whereas the current study focused only on racing-related SDFT injury. It has been found in previously published literature that the majority of tendon injuries occur during training rather than racing, therefore the lower prevalence reported here is expected (Ely et al., 2009).

The risk of SDFT injury in jump racing appeared to be significantly greater than flat racing. This finding has been identified previously (Stephen et al., 2003; Neundorf et al., 2010). The current study also identified an increased risk of SDFT injury for males compared with females, which has also previously been found (Kasashima et al., 2004; Perkins et al., 2005a; Oki et al., 2008; Kalisiak, 2012).

The number of career starts was found to be associated with SDFT injury in all horses as a single population, and also in jump racing horses; however, the direction of the association differed. Increasing numbers of starts was associated with an increase the risk of SDFT injury when all horses were considered together, but was associated with reduced risk in jump racing. This variable was not included in the final model of SDFT injury among flat racing horses. This finding may be an example of a 'healthy horse effect', where healthy horses are free to run more races, thus as the number of starts increases, the risk of SDFT injury reduces. Alternatively, this explanatory variable may be a measure of time spent at risk. Both flat racing and jump racing final SDFT injury models contained the career length variable, with increased odds of injury as the career length grew, suggesting again that increased time at risk was the more important variable.

In the model of SDFT injury for the whole BHA population, total winnings (in thousands of British pounds) was negatively associated with increased risk. Horses that accrue large amounts of winnings may be more robust and therefore able to race more often, and are possibly cared for more diligently than those horses that are not expected to accrue winnings. This variable was not included in the final models from flat or jump racing populations.

The significant interaction term between racing profile and sex in the final model of SDFT injury in the whole BHA population was included because more males than females compete in jump racing.

SDFT heritability is moderately to highly heritable in this population, and is of a magnitude that selective breeding could be used to reduce genetic risk in future. The models that produced the largest estimates of heritability in the current study were logistic sire models. Jump racing was found to be associated with increased risk of SDFT injury, although the heritability estimates of this outcome from the jump racing population are somewhat smaller than estimates from the flat racing population. This finding suggests that, although SDFT injuries are more likely to occur in jump racing, those SDFT injuries that are genetically predisposed do not occur more frequently in this population. A proportion of SDFT injuries may have been incidental, for example as a result of traumatic transection of the tendon during a collision or a fall, and these would not have contributed to the heritability of the condition.

Permanent environmental effects contributed substantially to the repeatability of SDFT injury. What these effects were is unknown, but they are hypothesized to be due to two causes. Healed tendon tissue displays markedly different mechanical properties to healthy tissue, and the degree of tendon 'scarring' may be influenced by environmental factors such as exercise during the healing period, nutritional plane etc. This alteration to the tendon tissue may, in part, constitute a permanent environmental effect, which would render the tendon susceptible to re-injury. Alternatively, some treatment methods may be employed that also increase the likelihood of repeated SDFT injury episodes. For example, firing of tendons is an antiquated, but still used 'treatment' for tendon injury. This intervention is used to initiate an acute healing process in a chronically damaged tendon, with the hope that this will lead to resolution of the injury. The efficacy of the procedure has not been proven, but it continues to be practiced in the UK. The resulting tendon damage may indeed lead to a 'healed' tendon, i.e. one that is devoid of active inflammation, but which has areas of scar tissue throughout the parenchyma. However, such tissue may be more susceptible to future damage.

4.3 General musculoskeletal conditions discussion

The importance of musculoskeletal conditions to horseracing cannot be understated. MSI are the most common medical problem affecting racehorses, and are prevalent in all jurisdictions, and across all race types (Williams et al., 2001; Stephen et al., 2003; Perkins et al., 2005b; Wilsher et al., 2006; Ramzan and Palmer, 2011; Bolwell et al., 2012). Financial and welfare implications of these conditions have led to a plethora of scientific articles exploring the factors involved in different conditions, but variation in case definitions, study power, and population characteristics mean that risk factor analyses are not usually comparable across studies. Here we present results of studies of the genetic risk of a number of important conditions in two racing jurisdictions. It is hoped that through genetic analyses, risk of these conditions could be decreased in future populations in conjunction with appropriate environmental modifications.

Tendon injury, or injury to the SDFT, was shown to be moderately to highly heritable for a binary trait in the current study. Oki et al. (2008) used different methods to those used here, but also found that SDFT injury was a heritable condition in the Japanese racing population (Oki et al., 2008). Results of SDFT injury analyses shown here suggest that greater care should be taken with older male jump racing horses, as their risk of SDFT injury appears to be greater than other horses'. Interestingly, the estimated heritability of SDFT injury was somewhat higher in horses that competed in flat races than those that competed in jump races, in Great Britain. Despite this, SDFT injury prevalence was greater in jump racing compared to flat racing. This suggests that environmental effects are acting to comparatively reduce SDFT injury incidence in flat racing, or increase it in jump racing, or both. The most logical explanation seems to be that jump racing affords more potential for idiosyncratic traumatic events such as falls or collisions between runners, which would increase SDFT injury prevalence without increasing heritability. Indeed, repeatability models of SDFT injury using records from flat racing or jump racing horses separately, showed that permanent environmental effects account for a larger proportion of repeatability in jump racing than in flat racing. Tendon injury heritability estimates were of sufficient magnitude that selective breeding based on EBVs as selection criteria could be useful in reducing future genetic risk.

Estimates of the heritability of OA and suspensory ligament injury were not always significantly different from zero. There were fewer cases of OA and suspensory ligament injury compared to tendon injuries in the HKJC data, which could have been one reason for the non-significant heritability estimates seen. Despite the increase in case numbers when OA and DJD were combined and modelled as 'joint disease', the heritability was still non-significant. In the case of suspensory ligament injury, sensitivity and specificity of diagnosis may have been low if the findings of clinical examinations were not always augmented by ultrasonography. Similarly, diagnosis of OA would only have been attempted in horses showing overt clinical signs, thus a proportion of subclinical cases would have been omitted, and without radiography, OA and OCD could not have been differentiated reliably. Rather than these conditions being lowly heritable, it is possible that these factors have acted to artificially reduce heritability estimates in the HK population (Bishop and Woolliams, 2010). Further studies incorporating diagnostic imaging in case definitions are warranted. However, much larger study sizes may be required if these injury definitions are very specific.

The EBVs generated from animal models of the conditions studied here were of poor accuracy overall. EBV accuracy is directly related to the speed of response to selection, thus selective breeding would only afford slow alterations in genetic risk, based on the data and case definitions used here. However, fracture was found to be genetically correlated with both OA and suspensory ligament injury in the HKJC dataset, indicating that reductions in risk of all of these conditions could be achieved by selecting based on one condition alone. The best condition to select on would be the one most easily and cheaply phenotyped, with the highest heritability, and the condition with the greatest EBV accuracies. Similarly, distal limb fracture and SDFT injury were found to be genetically correlated in the BHA dataset.

5. CHAPTER V. EXERCISE-INDUCED PULMONARY HAEMORRHAGE AND EPISTAXIS

5.1 Hong Kong EIPH/epistaxis analyses

In the HKJC data, there were 981 (19.4%) horses that were diagnosed with at least one episode of EIPH/epistaxis. The proportion of horses that were subjected to BAL and cytology for definitive diagnosis of EIPH was not known, therefore the case definition used here may include cases of both EIPH and epistaxis (either related to EIPH or due to other causes).

5.1.1 Multivariable analysis

Table 5-1 shows the results of likelihood ratio tests of multivariable sire and animal models of EIPH/epistaxis. Addition of a genetic random variable (sire or animal) significantly improved the fit of all models.

Table 5-1. Likelihood ratio test statistics and *p*-values of final models of EIPH/epistaxis. LRT statistic is determined by $2(logL_{full} - logL_{red})$ where *L* is likelihood, and *full* and *red* indicate the final model, and the 'reduced' final model, i.e. minus the genetic random variable only (sire or animal).

Model form		LRT statistic	*p*-value
Sire	Linear	58.10	<0.001
	Logistic	290.16	<0.001
Animal	Linear	85.12	<0.001
	Logistic	636.88	<0.001

Two fixed effects were found to be significantly associated with EIPH/epistaxis; hemisphere of origin, and career length (years) (Table 5-2). The odds of EIPH/epistaxis increased with increasing career length. Originating from the Southern hemisphere compared to the Northern hemisphere was found to be a risk factor for EIPH/epistaxis.

Table 5-2. Results of multivariable sire regression models of EIPH/epistaxis on linear and logistic scales.

		Linear model	Logistic model		
		Wald *p*-value	Wald *p*-value	Odds ratio	95% CI
Hemisphere	*1**	<0.001	<0.001	1 (REF)	
	2†			1.487	1.249 – 1.769
Career length (yrs)		0.008	0.009	1.056	1.015 – 1.098

*Northern hemisphere;† Southern hemisphere.

5.1.2 Heritability and genetic correlation estimation

Estimates of the heritability of EIPH/epistaxis, and their confidence intervals are shown in Table 5-3. The heritability of EIPH/epistaxis ranged from 0.13 to 0.33, and all estimates were significantly greater than zero. The largest estimate was produced using a logistic sire model, and the most precise estimate was produced using a logistic animal model.

Table 5-3. Heritability estimates of EIPH/epistaxis in the HKJC data.

Model form		h^2	Standard error	95% CI
Sire	Linear	0.2094	0.0422	0.1267 – 0.2921
	Logistic	0.3337	0.0741	0.1885 – 0.4789
Animal	Linear	0.1930	0.0361	0.1222 – 0.2638
	Logistic	0.1274	0.0270	0.0745 – 0.1803

Table 5-4 shows the results of bivariate analyses of EIPH/epistaxis with the named conditions. Genetic and phenotypic correlations are reported, with their standard errors. EIPH/epistaxis was found to be positively genetically correlated with OA (animal model), and negatively genetically correlated with tendon injury (animal model). Significant negative phenotypic correlations between EIPH/epistaxis and fracture, suspensory ligament injury and tendon injury were also found.

Table 5-4. Genetic and phenotypic correlations (with standard errors and 95% confidence intervals) between EIPH/epistaxis and the named conditions. Sire model results appear above animal model results, which are in bold type. P-values are based on LRT between the unconstrained model and the same model with the genetic correlation constrained to 0.00001.

	r_g*	s.e.†	95% CI‡	*p*- value	r_p§	s.e.	95% CI‡
Fracture	0.1784 **0.1430**	0.1963 **0.1906**	-0.2063 – 0.5631 **-0.2306 – 0.5166**	0.3537 **0.4386**	-0.0536 **-0.0540**	0.0141 **0.0142**	-0.0812 – -0.0260 **-0.0815 - -0.0259**
OA	0.2344 **0.2732**	0.1347 **0.1333**	-0.0296 – 0.4984 **0.0119 – 0.5345**	0.0773 **0.0368**	-0.0081 **-0.0106**	0.0143 **0.0144**	-0.0361 – 0.0199 **-0.0388 – 0.0176**
Suspensory ligament injury	-0.0139 **-0.3382**	0.2470 **0.2406**	-0.4980 – 0.4702 **-0.8098 – 0.1334**	1.0000 **0.1380**	-0.0301 **-0.0299**	0.0141 **0.0141**	-0.0577 - -0.0025 **-0.0575 - -0.0023**
Tendon injury	-0.2829 **-0.3574**	0.1632 **0.1394**	-0.6028 – 0.0370 **-0.6306 - -0.0842**	0.0919 **0.0145**	-0.1118 **-0.1123**	0.0140 **0.0142**	-0.1392 - -0.0844 **-0.1401 - -0.0845**

*Genetic correlation, †standard error, ‡confidence interval, §phenotypic correlation.

5.1.3 EIPH/epistaxis estimated breeding values

Figures 5-1 and 5-3 show mean EBV for EIPH/epistaxis, and mean EBV accuracy per year of birth cohort, based on linear and logistic models, respectively. EBV accuracies appeared to increase somewhat over the study period, but mean EBV per year of birth cohort showed no distinct trend. Regression of logistic EBVs on year of birth returned non-significant (*p*-values >0.05) values of the intercept and coefficient of year of birth. EIPH/epistaxis EBV accuracies ranged from 0.42 to 0.82 (mean 0.51) using the linear model, and from 0.18 to 0.78 (mean 0.36) using the logistic model. Spearman rank correlation coefficient for EBVs from linear and logistic models was 0.965 (*p*-value <0.001).

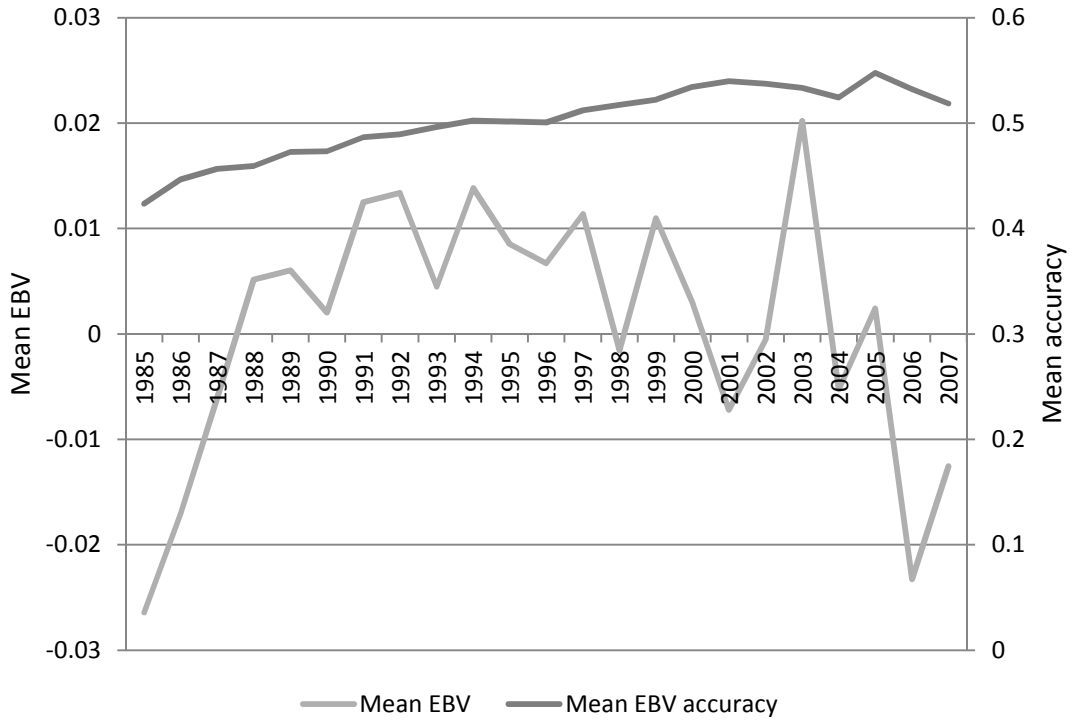

Figure 5-1. Mean EBV and mean EBV accuracies per year of birth cohort for EIPH/epistaxis, derived from the final linear animal model.

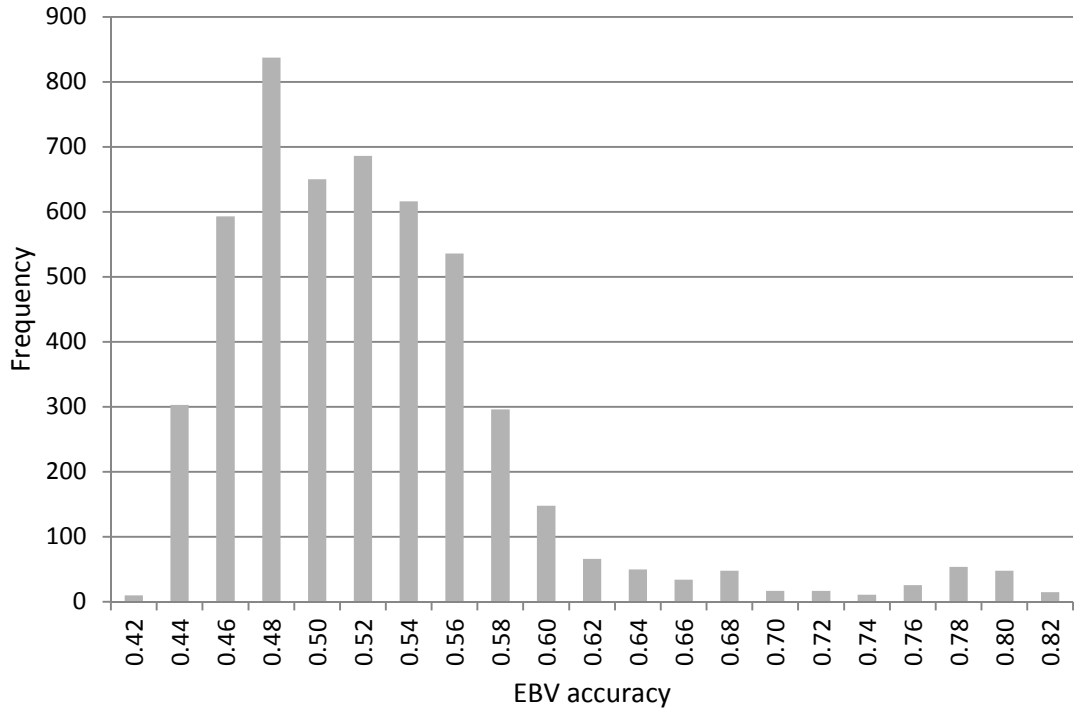

Figure 5-2. Frequency histogram of EBV accuracies for EIPH/epistaxis, derived from the final linear animal model.

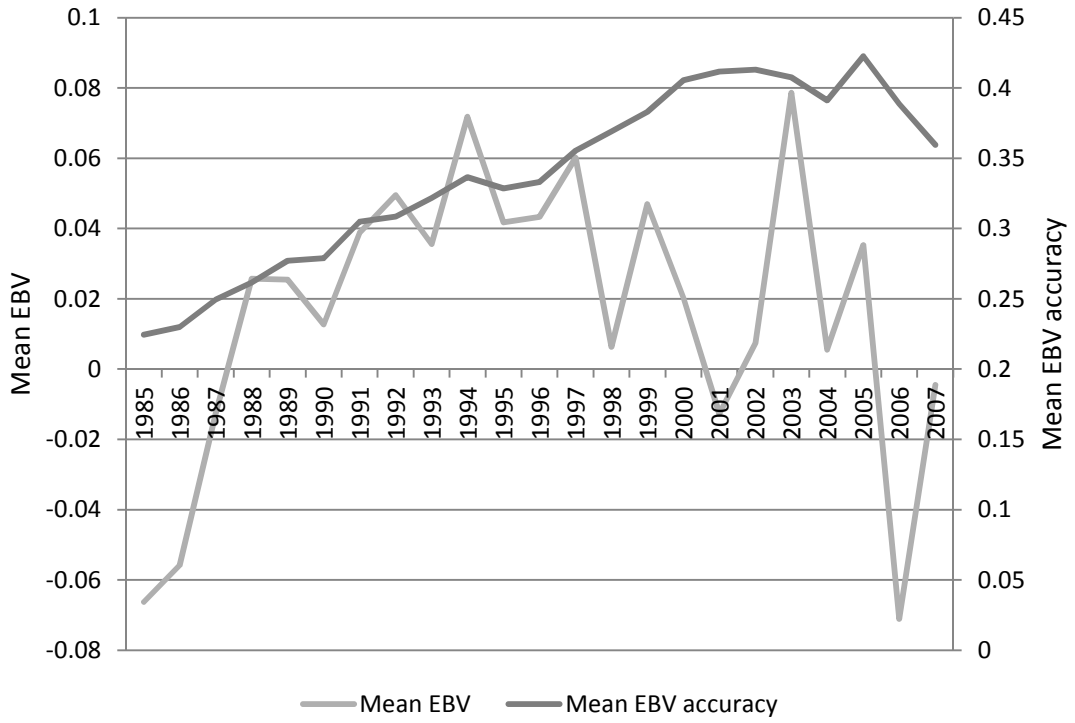

Figure 5-3. Mean EBV and mean EBV accuracies per year of birth cohort for EIPH/epistaxis, derived from the final logistic animal model.

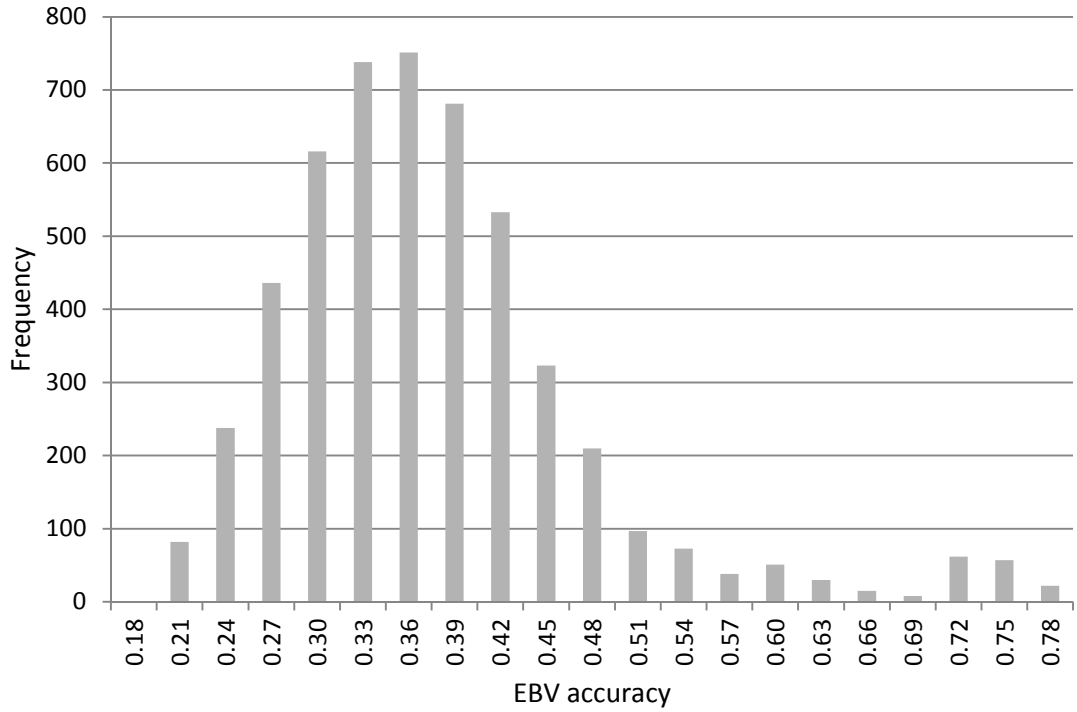

Figure 5-4. Frequency histogram of EBV accuracies for EIPH/epistaxis, derived from the final logistic animal model.

5.1.4 Hong Kong EIPH/epistaxis discussion

The prevalence of EIPH/epistaxis in this study was 19% overall. Most studies of EIPH focus either on epistaxis (using visual inspections of the nares only) or EIPH (using tracheobronchoscopy and/or BAL), but few combine these definitions. The reason for doing so in this study was that records of tracheobronchoscopic examinations, or BAL cytology were not available, thus diagnoses of 'bleeders' could have included both epistaxis and EIPH cases. The prevalence reported here is low compared to published EIPH prevalence, but high compared to previously published epistaxis prevalence (Takahashi et al., 2001; Hinchcliff et al., 2009).

Fixed effects found to be associated with EIPH/epistaxis in this study were the hemisphere of origin and career length. The variable of hemisphere of origin may describe different genetic groups of horses, or different management practices prior to import into Hong Kong that have long lasting effects on the risk of EIPH/epistaxis. The odds of EIPH/epistaxis were found to increase with increasing career lengths. Three recorded episodes of epistaxis necessitate compulsory retirement for Thoroughbreds in Hong Kong, therefore horses that are 'bleeders' might be expected to have shorter career lengths, although this is at odds with the association detected. An alternative explanation for this relationship would be that the lungs or other parts of the respiratory system of older horses are more susceptible to EIPH/epistaxis for whatever reason, or the training of young horses may differ from the training of older animals, leading to increased risk in the latter. Longer careers have also been found to increase the risk of epistaxis in other studies (Newton et al., 2005), and these temporal variables have been posited to be proxies for 'time spent racing', leading to increased likelihood of EIPH/epistaxis as a type of 'repetitive strain injury' to lung parenchyma.

The heritability of EIPH/epistaxis in this study was high for a binary trait, at 0.13 to 0.33. These estimates are comparable to Weideman et al. (2004), who reported heritability of epistaxis (assumed to be due to EIPH) to be 0.40 using a logistic sire model, and 0.23 using a logistic animal model. In both the current study, and the study by Weideman et al. (2004), the largest estimates were achieved using sire logistic models. Due to the magnitude of the heritability of this condition, selective breeding could be effective in inciting change in genetic risk. EBVs for this condition were found to be moderately

accurate, thus the speed of genetic gain could be favorable should such breeding strategies be widely implemented.

EIPH/epistaxis was found to be positively genetically correlated with OA, using an animal model. Perhaps genes that encode common inflammatory pathways exhibit pleiotropy, leading to increased frailty of pulmonary capillaries following exercise-induced insult, and also to inflammation of joint tissues. Molecular genetic studies would be required to identify the nature of this association. It is interesting to note that EIPH/epistaxis was negatively genetically correlated with tendon injury in this study. It is not clear why genes that confer risk of EIPH/epistaxis might also be protective against tendon injury, and this relationship may instead be due to the genes involved being sited close together in the equine genome, but again a detailed analysis of the genes involved is warranted to identify the underlying nature of this correlation.

5.2 BHA epistaxis analysis results

5.2.1 Results from analyses using all horses

There were 1667 horses that were diagnosed at least once with epistaxis in the 'per horse' (PH) dataset (n=78151, 2.1%). In the 'repeated records' (REP) dataset, there were 1960 cases of epistaxis recorded, giving an incidence rate of 2.2 cases per 1000 starts (n=900757). When only horses that competed in at least one flat race were considered, the incidence rate of epistaxis was 1.5 cases per 1000 starts (1036/670322), and when only horses that competed in at least one jump race were considered, the incidence rate was 2.9 cases per 1000 starts (1411/490659).

5.2.1.1 PH dataset multivariable analysis

Results of multivariable modeling of epistaxis in the BHA PH dataset are shown below. Table 5-5 shows likelihood ratio test results of epistaxis final models. All models were significantly improved by the addition of a genetic random variable.

Table 5-5. Likelihood ratio test statistics and *p*-values of final models of epistaxis. LRT statistic is determined by $2(logL_{full} - logL_{red})$ where *L* is likelihood, and *full* and *red* indicate the final model, and the 'reduced' final model, i.e. minus the genetic random variable only (sire or animal).

Model form		LRT statistic	*p*-value
Sire	Linear	56.64	<0.001
	Logistic	120.66	<0.001
Animal	Linear	70.96	<0.001
	Logistic	674.10	<0.001

Table 5-6 shows the significant associations of four fixed effects with epistaxis, and an interaction between career length and racing profile. The odds of epistaxis were found to increase with increasing career length (years), increasing number of starts, and with jump or mixed race type horses compared with exclusively flat racing horses. Being female was found to be protective.

Table 5-6. Results of multivariable sire regression models of epistaxis on linear and logistic scales.

		Linear model	Logistic model		
		Wald *p*-value	Wald *p*-value	Odds ratio	95% CI
Career length (yrs)		<0.001	<0.001	1.441	1.352 – 1.536
Number of starts		<0.001	<0.001	1.011	1.007 - 1.016
Racing profile	*1**	<0.001	<0.001	1 (REF)	
	2†			2.534	2.076 - 3.092
	3‡			3.510	2.725 – 4.522
Sex	Male	<0.001	<0.001	1 (REF)	
	Female			0.777	0.680 – 0.886
Career length*profile *2*		<0.001	<0.001	0.908	0.855 - 0.963
Career length*profile *3*		<0.001	<0.001	0.766	0.717 – 0.817

*Horses that competed exclusively in flat racing; †Horses that competed exclusively in jump racing; ‡Horses that competed in both flat and jump racing.

5.2.1.2 Heritability and genetic correlation estimation

Heritability estimates of epistaxis and their confidence intervals are shown in Table 5-7. All heritability estimates were significantly greater than zero. Estimates ranged from 0.02 to 0.26, and the sire logistic model provided the largest estimate.

Table 5-7. Heritability estimates of epistaxis in the BHA PH data.

Model form		h^2	Standard error	95% CI
Sire	Linear	0.0226	0.0034	0.0159 – 0.0293
	Logistic	0.2625	0.0430	0.1783 – 0.3467
Animal	Linear	0.0238	0.0034	0.0171 – 0.0305
	Logistic	0.1698	0.0253	0.1203 – 0.2194

Epistaxis was found to be positively genetically correlated with SDFT injury (Table 5-8), but no other genetic or phenotypic correlations were significant.

Table 5-8. Genetic and phenotypic correlations (with standard errors and 95% confidence intervals) between SDFT injury, distal limb fracture and epistaxis. Sire model results appear above animal model results, which are in bold type. P-values are based on LRT between the unconstrained model and the same model with the genetic correlation constrained to 0.00001.

	r_g*	s.e.†	95% CI‡	*p*- value	r_p§	s.e.	95% CI‡
Distal limb fracture	0.1958 **0.1753**	0.1221 **0.1166**	-0.0435 – 0.4351 **-0.0532 – 0.4038**	0.7083 **0.2401**	0.0004 **0.0003**	0.0036 **0.0036**	-0.0067 – 0.0075 **-0.0068 – 0.0074**
SDFT injury	0.2942 **0.3021**	0.0748 **0.0697**	0.1476 – 0.4408 **0.1655 – 0.4387**	0.0012 **0.0001**	0.0067 **0.0066**	0.0036 **0.0036**	-0.0004 – 0.0138 **-0.0005 – 0.0137**

*Genetic correlation, †standard error, ‡confidence interval, §phenotypic correlation.

5.2.1.3 Estimated breeding values for epistaxis

Animal models of epistaxis were used to extract EBVs, and mean values per year of birth cohort are shown in Figures 5-5 and 5-7. The mean accuracies of these EBVs were distributed as shown in Figures 5-6 and 5-8. Mean EBV appeared to become more negative through the twenty-first century, and accuracies tended to improve until the last few years of the study period. Regression of logistic EBVs on year of birth was equivalent to a decrease of -0.0070 ± 0.0002 grades per annum, indicating a small but significant decreasing (improving) trend. EBV accuracies ranged from 0.08 to 0.90 (mean 0.27) from linear models, and from 0.05 to 0.68 (mean 0.28) in logistic models. Spearman rank

correlation coefficient between EBVs generated by these two model forms was 0.991 (*p*-value <0.001).

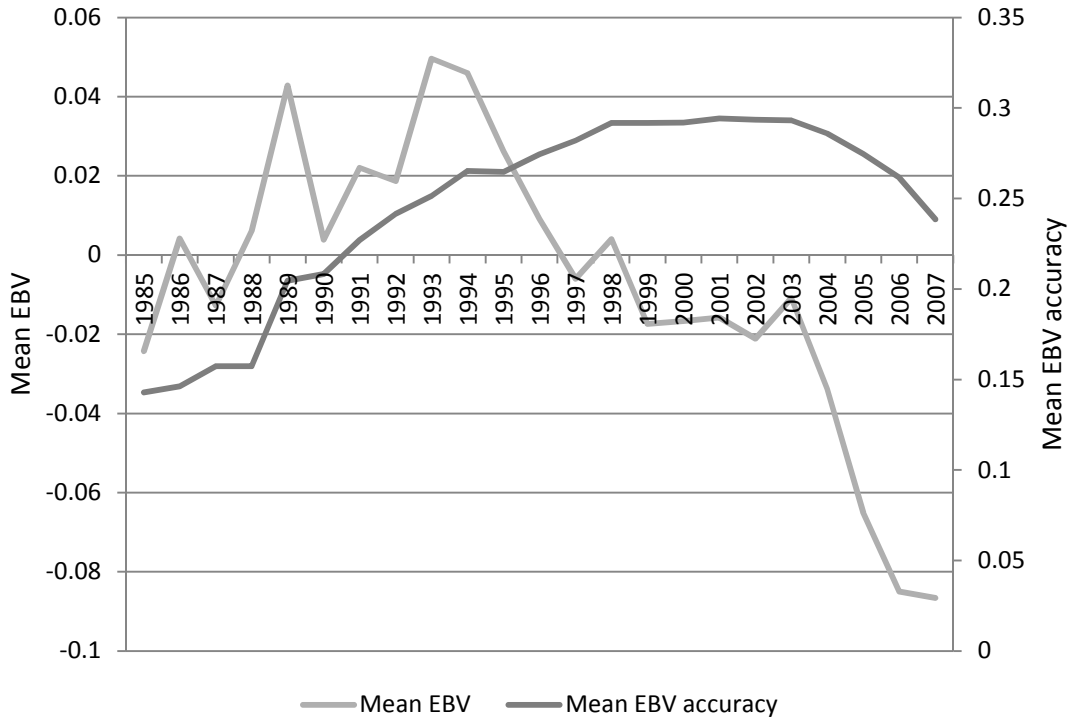

Figure 5-5. Mean EBV and mean EBV accuracies per year of birth cohort for epistaxis, derived from the final linear animal model.

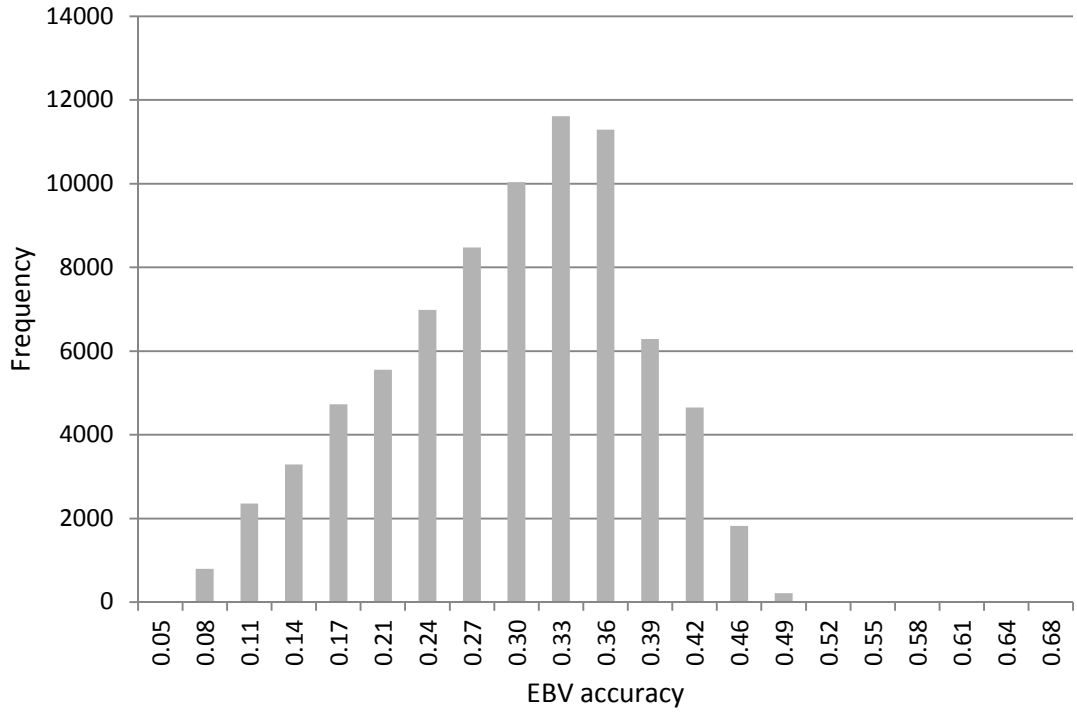

Figure 5-6. Frequency histogram of EBV accuracies for epistaxis, derived from the final linear animal model.

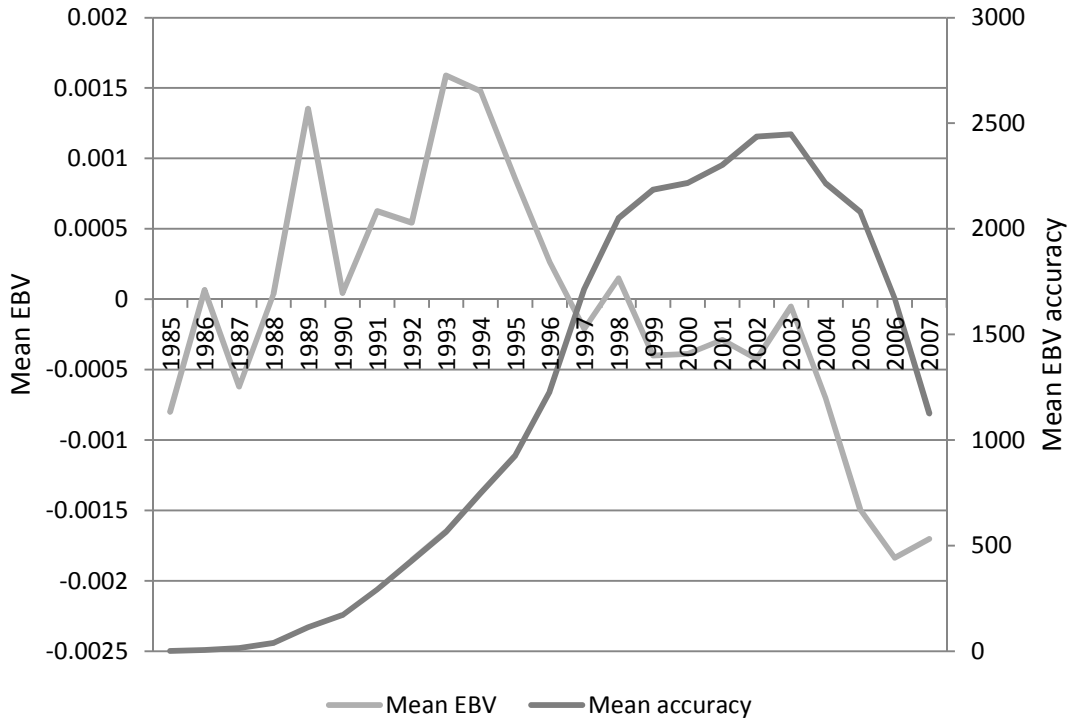

Figure 5-7. Mean EBV and mean EBV accuracies per year of birth cohort for epistaxis, derived from the final logistic animal model.

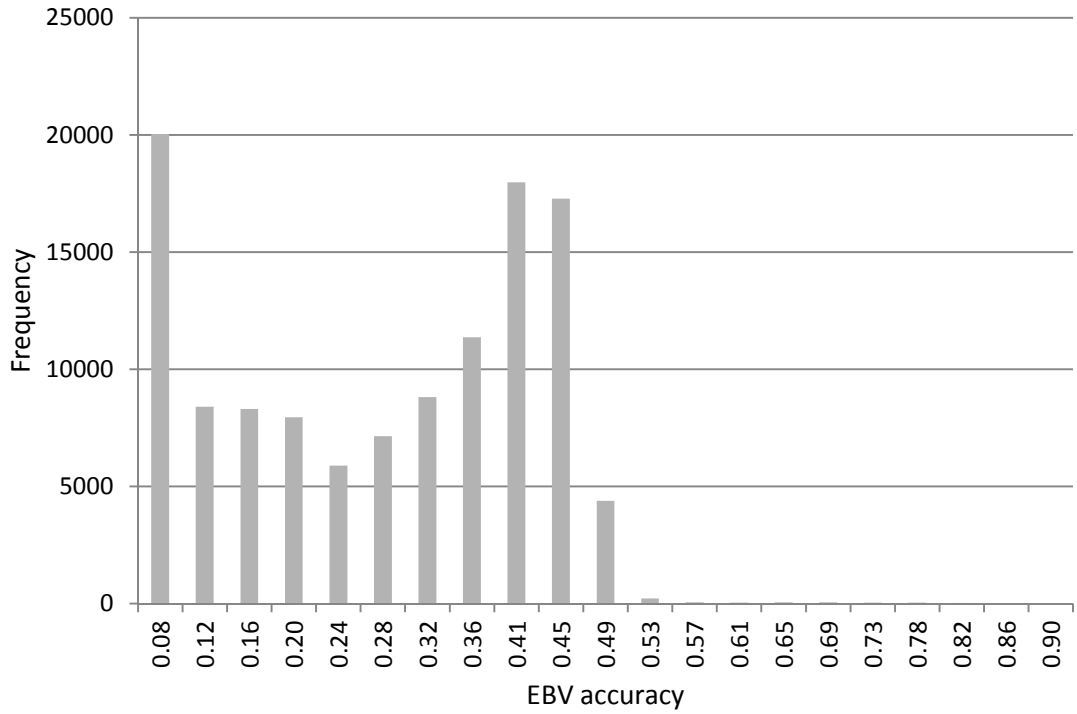

Figure 5-8. Frequency histogram of EBV accuracies for epistaxis, derived from the final logistic animal model.

5.2.2 BHA REP dataset heritability and repeatability

The BHA REP dataset was used to estimate heritability, repeatability and permanent environmental variance (the effects of environmental influences that last for the lifetime of the horse) of epistaxis in the UK, using linear sire and animal models (Table 5-9). Permanent environmental variance was negligible in the animal model, where heritability accounted for all of the repeatability of the condition.

Table 5-9. Heritability, permanent environmental variance and repeatability of epistaxis in the BHA REP dataset.

	h^2* (s.e.)	σ^2_{pe}† (s.e.)	r≠ (s.e.)
Sire	0.0020 (0.0003)	0.0058 (0.0002)	0.0053 (0.0002)
Animal	0.0017 (0.0003)	0.0000 (0.0000)	0.0017 (0.0003)

*Heritability; † permanent environmental variance;‡ repeatability.

5.2.3 Flat racing population

Within the population of horses that competed in any flat races in the BHA data, 865 horses were diagnosed at least once with epistaxis (1.7%, n=49660).

5.2.3.1 Multivariate analysis

Significant improvement in model fit was seen in all models of the flat racing population, when sire or animal were added as random effects (Table 5-10).

Table 5-10. Likelihood ratio test statistics and *p*-values of final models for epistaxis in flat racing horses. LRT statistic is determined by $2(logL_{full} - logL_{red})$ where *L* is likelihood, and *full* and *red* indicate the final model, and the 'reduced' final model, i.e. minus the genetic random variable only (sire or animal).

Model form		LRT statistic	*p*-value
Sire	Linear	47.46	<0.001
	Logistic	5940.16	<0.001
Animal	Linear	122.62	<0.001
	Logistic	9579.27	<0.001

Length of racing career, number of starts and sex were found to be associated with epistaxis in the flat racing population. Both increased career length and increased number of starts were associated with increased odds, and again female sex was associated with reduced odds of epistaxis (Table 5-11).

Table 5-11. Results of multivariable sire regression models of epistaxis in flat racing horses on linear and logistic scales.

		Linear model	Logistic model		
		Wald *p*-value	Wald *p*-value	Odds ratio	95% CI
Career length (yrs)		<0.001	<0.001	1.317	1.248 – 1.389
Number of starts		<0.001	<0.001	1.010	1.005 – 1.014
Sex	Male	<0.001	<0.001	1 (REF)	
	Female			0.539	0.418 - 0.695

Heritability estimates of epistaxis in the flat racing population ranged from 0.02 to 0.28 (Table 5-12). The logistic sire model produced the largest estimate, and the two linear estimates were almost identical.

Table 5-12. Heritability estimates of epistaxis in flat racing horses in the BHA data.

Model form		h^2	Standard error	95% CI
Sire	Linear	0.0180	0.0040	0.0103 – 0.0252
	Logistic	0.2763	0.0614	0.1560 – 0.3967
Animal	Linear	0.0180	0.0038	0.0106 – 0.0255
	Logistic	0.1761	0.0362	0.1051 – 0.2471

5.2.4 Jump racing population

Within the population of horses that competed in any jump races in the BHA data, 1204 horses were diagnosed at least once with epistaxis (3.0%, n=40135).

5.2.4.1 Multivariable analysis

Within the jump racing BHA population, all models were significantly improved by addition of a genetic random variable (Table 5-13).

Table 5-13. Likelihood ratio test statistics and *p*-values of final models for epistaxis in jump racing horses. LRT statistic is determined by $2(logL_{full} - logL_{red})$ where *L* is likelihood, and *full* and *red* indicate the final model, and the 'reduced' final model, i.e. minus the genetic random variable only (sire or animal).

Model form		LRT statistic	*p*-value
Sire	Linear	68.72	<0.001
	Logistic	4771.36	<0.001
Animal	Linear	122.62	<0.001
	Logistic	9579.27	<0.001

Length of career, number of starts and sex were found to be significantly associated with epistaxis in the jump racing population. These associations were similar to those seen in the flat racing population; however, there was also a significant interaction between career length and number of starts in the jump racing final models (Table 5-14).

Table 5-14. Results of multivariable sire regression models of epistaxis in jump racing horses on linear and logistic scales.

		Linear model	Logistic model		
		Wald *p*-value	Wald *p*-value	Odds ratio	95% CI
Career length (yrs)		<0.001	<0.001	1.359	1.296 – 1.425
Number of starts		<0.001	<0.001	1.048	1.038 – 1.059
Sex	Male	0.003	<0.001	1 (REF)	
	Female			0.759	0.643 - 0.896
Interaction*		<0.001	<0.001	0.993	0.991 – 0.994

*First order interaction between career length (years) and number of starts.

Epistaxis heritability estimates among jump racing horses ranged from 0.03 to 0.26. Again the logistic sire model gave the largest estimate, and both animal and sire model linear estimates were similar (Table 5-15).

Table 5-15. Heritability estimates of epistaxis in jump racing horses in the BHA data.

Model form		h^2	Standard error	95% CI
Sire	Linear	0.0310	0.0060	0.0197 – 0.0429
	Logistic	0.2640	0.0531	0.1599 – 0.3681
Animal	Linear	0.0309	0.0056	0.0199 – 0.0418
	Logistic	0.1498	0.0296	0.0919 – 0.2078

5.2.5 BHA epistaxis discussion

The prevalence of epistaxis in this study was broadly similar to published figures (Williams et al., 2001; Weideman et al., 2004; Hinchcliff et al., 2009). A total of 1667 horses were diagnosed at least once with epistaxis (2.1%). This figure was 1.7% when only flat racing horses were considered (865/49660), and was 3.0% when only jump racing horses were considered (1204/40135). Among all race types, the incidence rate of epistaxis was 2.2 cases per 1000 starts (1960/900757). When only starts from horses that competed in flat races were considered, there were 1.5 cases of epistaxis per 1000 starts (1036/670322), and when only starts from horses that competed in jump races were considered, there were 2.9 cases per 1000 starts (1411/490659).

The heritability of epistaxis was found to be similar among flat racing and jump racing horses, and when all horses were considered together. These estimates were lower than those reported by Weideman et al. (2004), who reported an estimated heritability of 0.23 to 0.4. Nevertheless, these estimates indicate a significant genetic component to the risk of epistaxis in this population, which appears to be similar across different types of horse with different racing profiles. Logistic sire models were consistently shown to produce the largest estimates of heritability.

The length of racing career and number of starts were associated with epistaxis risk in all models. A similar association between career length and epistaxis has been identified previously (Newton et al., 2005). Other studies have identified an association between horse age and epistaxis, and this could be a reason behind both the association of epistaxis with career length, and with the number of starts (Takahashi et al., 2001; Watkins and Stewart, 2008). Jump racing compared to flat racing has been shown to be a risk factor for epistaxis, and this was also seen in this study, as the prevalence and incidence rates of epistaxis were greater for jump racing compared to flat racing (Takahashi et al., 2001; Newton et al., 2005).

Within the jump racing BHA population, an interaction term between career length and the number of starts was significant. This interaction was not significant in the flat racing population. In jump racing, this interaction suggests that career length and number of starts are both important in epistaxis risk, but they also influence each other, as may be expected (i.e. longer careers contain more starts, and horses that start more races have longer careers). However, in flat racing, career length and the number of starts were not shown to interact with each other, suggesting that flat race horses could have long careers with few starts, or shorter, more intense racing careers, but that no consistent relationship exists between these variables.

Results of repeatability models in the current study showed that permanent environmental effects are not important to the repeatability of epistaxis in this population. Heritability based on repeatability models was very low at around 0.002. Due to the large difference in heritability estimates between repeatability and all other model types, it is suggested that these models are not appropriate for use with binary data of this nature.

5.3 General EIPH/epistaxis discussion

Epistaxis is often assumed to be caused by haemorrhage of pulmonary origin, and is used as a convenient marker of serious EIPH in risk factor analyses (Takahashi et al., 2001; Weideman et al., 2004; Newton et al., 2005). Unfortunately, these studies do not often confirm that all cases of epistaxis occur as a result of EIPH. A proportion of cases of epistaxis may be caused by alternative problems such as head trauma, which may reduce the number and strength of associations found. Risk factors found to be associated with epistaxis include increasing age, female sex, shorter race distances, harder going and racing over jumps (Takahashi et al., 2001; Newton et al., 2005). Risk factor analyses of EIPH have failed to demonstrate similar associations, therefore it is unknown whether risk factors for EIPH and epistaxis differ (Hinchcliff et al., 2010). It is possible that factors that increase the risk of pulmonary haemorrhage are very different to those of epistaxis, and epistaxis-related risk factors reflect the risk of blood travelling from the lungs to the nostrils (and not being swallowed), rather than true pulmonary haemorrhage. Epistaxis has been found to have a significant genetic component, but the genetics of EIPH specifically have not been studied. It may therefore be the case that EIPH itself is not heritable, but factors such as anatomical variations in pulmonary tissues and the propensity for EIPH to be of sufficient seriousness to cause epistaxis mean that epistaxis is heritable. This question cannot be answered until genetic studies of EIPH are conducted.

Previous studies have identified jump racing as a risk factor for epistaxis, which was also found in this study (Takahashi et al., 2001; Williams et al., 2001; Newton et al., 2005). In the UK data in the current study, some evidence was found that jump racing horses were a different genetic group to flat racing horses, thus this risk factor may be partly genetic in origin. Jump racing may also provide added risks of epistaxis through the increased forces generated through the limbs, and transmitted to the lungs (Schroter et al., 1998). This impact-induced haemorrhage theory was posited by Schroter et al. (1998), and although a link between locomotor impacts and EIPH was found, this increase in risk of EIPH may subsequently increase the risk of epistaxis. This theory of EIPH pathogenesis may be supported by the finding that running on harder ground is associated with increased risk, although this risk factor was not available for the current study (Newton, 2008). The action of jumping may also favour movement of blood from lung tissue towards the nostrils, and jumping may be associated with greater effort, thus be more likely to result in haemorrhage through capillary stress failure (West and Mathieu-Costello, 1994). The

prevalence of epistaxis was higher in jump racing horses compared with flat racing horses in the current study, but heritability estimates of epistaxis were similar in both groups. This suggests that jump racing exposes horses to environmental risk factors that increase the likelihood of epistaxis.

The only significant risk factor found to be associated with EIPH (of all levels of severity) in veterinary literature was colder air temperature by Hinchcliff et al. (2010), using a cohort of Australian flat racing Thoroughbreds. This factor has not been found to be associated with epistaxis in any studies. The reasons behind this association are unknown, but the authors suggested that there was a link between cold air inhalation at exercise and pulmonary inflammation (Hinchcliff et al., 2010). EIPH of more serious grades was found to be associated with increased numbers of starts over a lifetime, with races of moderately long distance, and with increased winnings.

An opposite direction of association between horse sex and epistaxis risk was found in this study compared to previous studies (Takahashi et al., 2001). This study found that males were more at risk than females, whereas Takahashi et al. (2001) found that females were more at risk compared with sexually intact males. Sex could not be adequately examined in the HKJC data due to the small number of female horses in the data. The reason for the discrepancy found is unclear. No association between EIPH and horse sex was found by Hinchcliff et al. (2010).

Increased age was found to be a risk factor for epistaxis in previous studies (Takahashi et al., 2001; Williams et al., 2001). Both career length and number of starts in the current study could reflect a similar accumulation of time at risk, or of ageing change, but age itself was not found to be independently associated with epistaxis risk. Hinchcliff et al. (2010) found that an increased number of starts over the lifetime of the horse was associated with increased risk of EIPH of grades above or equal to 1. These findings collectively suggest that older horses that have competed in more races are predisposed to EIPH and subsequent epistaxis, suggesting either an accumulation of damage to the tissues involved, age-related susceptibility, or simply that time at risk is important in this condition. It has been proposed in previous studies that age, career length, number of starts and other time-related variables are proxies for ‘time spent racing’ (Binns et al., 2012), and that EIPH is a type of ‘repetitive-strain injury’ to lung parenchyma. The nature of this accumulated lung damage has not been fully elucidated and validated, however one study

found an association between EIPH and veno-occlusive lesions in areas of lung tissue where bleeding and remodeling were evident (Williams et al., 2008). This histological alteration provides one explanation for an increased likelihood of EIPH/epistaxis following increased time spent racing (or exercising).

EIPH has long been thought to affect horse health negatively. Some studies have recorded racehorse deaths attributable to severe pulmonary haemorrhage (Johnson et al., 1994; McKee, 1995; Lyle et al., 2012). The number of deaths recorded is extremely low when the very large prevalence of EIPH is taken into account. It is unknown whether these horses reflected a separate, more extreme population of EIPH sufferers with their own set of risk factors. Studies of the effects of milder EIPH, or of epistaxis, on horse health are lacking. Poor performance has been linked to both epistaxis and serious cases of EIPH, and this link has fuelled much of the EIPH/epistaxis research that has been conducted to date. Further genetic studies of EIPH or epistaxis should be very careful to identify their breeding goals, and be clear as to whether the ultimate goal is to improve and safeguard horse health and welfare, or to improve performance. If welfare is not affected by epistaxis, and it is purely a cosmetic, public relations or performance issue, veterinary research should only proceed with caution, to maintain the integrity of the profession.

Interestingly, epistaxis was positively genetically correlated with SDFT injury in the BHA data of this study, but EIPH/epistaxis was negatively genetically correlated with tendon injury in the HKJC data. Obviously, these conditions are not defined identically in the HKJC and BHA data, thus they are likely to reflect slightly different spectra of conditions. Despite this, the different direction of correlation is notable. It seems unlikely that the difference between epistaxis at the race course and EIPH/epistaxis in racing or training, or between SDFT injury in racing and tendon injury in racing or training, are significant enough to produce this discrepancy in correlations. An alternative explanation is that Hong Kong and UK Thoroughbred populations are different genetic groups, with different genetic bases to these conditions. This may mean that the genes or even genetic aetiologies involved in epistaxis risk vary in different populations. Molecular genetic studies could be undertaken to ascertain whether this is the case.

This study provides further evidence to Weideman et al. (2005) that epistaxis is a heritable condition in racing Thoroughbreds. An unknown proportion of the population of Thoroughbreds in Hong Kong may have been confirmed as ‘bleeders’ through the use of

tracheobronchoscopy, thus the case definition in this part of the study was EIPH/epistaxis. Whether the pathogenesis of true EIPH involves stress failure of capillaries under high pressures, locomotory-impact induced trauma, impaired innate immunity or a combination of these factors, the involvement of genetic risk cannot be stated with confidence based on existing research (West and Mathieu-Costello, 1994; Schroter et al., 1998; Michelotto et al., 2011). Epistaxis as related to EIPH is heritable, but whether this heritability reflects the underlying heritability of EIPH, or of factors which make blood travel more swiftly and easily to the nostrils from the alveoli, is unclear, and bears further scrutiny.

6. CHAPTER VI. PERFORMANCE AND LONGEVITY

6.1 A review of literature

Horseracing is a commercial enterprise that relies upon the provision of healthy animals, capable of competing at their maximum level of effort. Naturally, there is turnover of the horse population due to old age, unavoidable health concerns or managemental decisions. Poor racing performance is also a significant cause of wastage of horses. Trainers and horse owners are interested in retaining the good quality animals they have by preventing health problems, but are also bound to be motivated to accrue the best performing animals possible. Research has been conducted in many racing jurisdictions to attempt to identify genetic parameters associated with performance, in the hope that careful mating decisions, and/or genetic screening tests could identify those animals likely to be capable of top class performances in later life. A number of molecular genetic studies have identified SNPs associated with racing performance, the most notable of which was the g. 66493737C>T on ECA (*Equus caballus* chromosome)18, close to the *MSTN* (myostatin) gene, but many others have been identified (Gu et al., 2009; McGivney et al., 2009; Hill et al., 2010; Schroder et al., 2011; Hill et al., 2012; Tozaki et al., 2012). Typically, these SNPs are close to genes associated with muscle strength-related features, fat oxidation, and mitochondrial actions. Understanding the mechanisms behind the genes identified will provide insights into the nature of elite physical performance. Since 2010, some commercial genetic tests have become available that can genotype individuals for panels of exercise-related SNPs, or single SNPs such as g. 66493737C>T (www.equinome.com). Interpretation of these test results for individual horses, or as part of breeding programs, is of paramount importance, as eventual phenotypes depend on many more genes than are interrogated as part of these tests, as well as significant environmental influences. Client education is therefore essential if the use of such tests for multifactorial traits is to become widespread. An understanding of the likely proportion of variance in the performance phenotype associated with individual or groups of SNPs is essential to managing the expectations of horse breeders that intend to use such genetic tests.

Other studies of equine performance genetics have focused on the heritability of different racing-related traits as a whole. Many of these studies used racing time as the dependent variable in their analyses, in the hope that it is the most appropriate measure of performance, taking the effects of competitors into account (Mota et al., 2005; Ekiz and Kocak, 2007; Buxadera and da Mota, 2008; Bakhtiari and Kashan, 2009; Park, 2011). Heritability estimates of racing time have been consistently reported to be between 0.04

and 0.35, with repeatability of racing time estimated to be up to 0.59 (Mota et al., 2005; Ekiz and Kocak, 2007; Buxadera and da Mota, 2008; Bakhtiari and Kashan, 2009; Park, 2011). The heritability of racing time was found to decrease with increasing race distance in all of these studies, suggesting that longer races give more time for environmental effects to play a part. The effects of the contemporary group and of the jockey were found to contribute significantly to racing time in one study, and the variation attributable to the jockey increased as the race distance increased (Park, 2011). Racing time is a relatively easy variable to measure, and can be 'calculated' for non-placed horses based on distance behind the winner, for races where race time is not routinely recorded for horses coming fourth or later (Mota et al., 2005). Significant sources of variation in racing time include the going, jockey, starting position, age and sex of the horse (Mota et al., 2005; Buxadera and da Mota, 2008; Bakhtiari and Kashan, 2009; Park, 2011). This variable has the advantage of being a repeated measure on the same individual, thus repeatability models can be used to assess inter- and intra-animal variation.

Studies of horse performance have also used earnings, intensity of racing, and coefficients of racing success as dependent variables in heritability analyses (Sobczynska, 2010; Tozaki et al., 2012). Tozaki et al. (2012) studied the lifetime earnings of Japanese racehorses, and found that the heritability of this trait was between 0.11 and 0.25 for linear and non-linear models, respectively. In a study of Polish Arabian racehorses, Sobczynska et al. (2010) used coefficients of success (CS) and intensity (CI) of racing, which are used to rank horses currently. The CS was defined as the horse's annual earnings divided by the mean earnings of horses of the same age and sex, competing in the same year. CI was defined as the number of starts a horse completed in a given season, compared to the mean number of starts by horses of the same sex in that season. Both coefficients were investigated for horses of three or four years of age only. Heritability estimates for CI at ages three and four years were low, whereas estimates of the heritability of CS were higher, at 0.41 and 0.18 for three and four year olds, respectively. The number of races a horse runs per season is partly determined by its health, but is also largely a factor of human decision-making. In this study, CS did not take into account the number of starts per horse, thus earnings was a crude measure of performance, which nevertheless was moderately heritable (Sobczynska, 2010).

These studies have shown that performance (based on racing times or earnings-related factors) is a heritable trait, thus we may be able to improve the overall performance of a

population of Thoroughbreds through careful selective breeding based on overt phenotypes, with or without the use of genotyping tests. It is hoped that this improvement could reduce the wastage of horses from racing in the longer term, and might be a proxy measure for optimized horse health, since only the fittest and most healthy animals will be capable of good performances. Selective breeding of those horses with the most perceived 'talent' has been performed since the sport of horse racing was conceived, and continues today. Choices on which horses are to be crossed are predominantly made without consideration of the results of genetic tests. It has been noted elsewhere that racehorse performances have improved little if at all over the last century, despite this selective breeding (Gianola, 1982; Verheyen and Wood, 2004). It is thought that this apparent lack of improvement is partly due to the use of dependent variables that do not accurately reflect performance, or that are indirect measures at best. If there has truly been no improvement in performance over this time, this may be because a horse's performance is affected by a large number of different, occasionally antagonistic variables such as musculature, weight, long bone length or ability of competitors. Selective breeding that favours only some of these aspects may lead to negative effects on others, thus producing no discernible difference in overall racing performance. Future genetic studies designed to investigate the most important factors in racing performance would bring some clarity to this problem, and may allow for faster genetic gain where more arbitrary choices of matings have failed to lead to performance improvements over past generations.

Although every trainer and racehorse owner has an interest in winning races, the quality of events varies greatly, and a proportion of stakeholders can be profitable by concentrating their efforts on producing consistent winners in lower-ranked races, rather than striving for elite-level animals. The durability of horses at all levels of racing is, however, important, as maximal financial gains can only be made where each animal is robust enough to run a number of races, with a good chance of winning, unless the horse is of such quality that it is capable of winning a small number of very lucrative races. The number of races a horse may run over its career is heavily affected by managemental decisions, but is in part determined by that animal's athletic ability, temperament, and health. If durability (i.e. the number of starts in a career) were the breeding goal, it is possible that both performance and horse health would be improved concurrently, to the benefit of owners, veterinarians, and most importantly, horse welfare. Investigations of the heritability of measures of performance and durability must include genetic correlations with health traits, to ensure that breeding goals would not inadvertently increase the risk of adverse health events.

To investigate measures of performance and durability in the current datasets, and their relationship with various health concerns, a number of traits were subjected to heritability analyses in HK and BHA data; length of career (years), number of career starts, and career earnings. The methods and results of these analyses follow.

6.2 Horse durability

6.2.1 Career length

6.2.1.1 HKJC career length heritability analyses

6.2.1.1.1 Results

Length of racing career in years was defined as the number of days between the first race run and the date of retirement, divided by 365. A career length of zero indicated horses that were recorded as retiring on their first race day, or those that did not race. This variable was distributed as shown in Figure 6-1. A Shapiro-Wilks test of normality revealed a significant departure from the Normal distribution (p-value <0.05), therefore the variable was log-transformed after adding one day to the length of each career (to avoid generating log(0)). This derived dependent variable will be referred to as CL. Model building proceeded as described previously (see Chapter 2), using linear animal models of CL and testing potential fixed effects that were used previously.

The final linear animal model of CL included the fixed effect of the year of birth (where 1985 to 1987 were grouped, years between 1988 and 2005 were individually coded, and 2006 and 2007 were grouped), which was significant at p-value <0.001. Addition of animal as a random variable, where the non-independence of animals was defined through the pedigree, significantly improved the fit of the final model (p-value <0.001). Heritability of CL was estimated to be 0.2167 (s.e. 0.0400, 95% confidence intervals 0.1383 to 0.2951).

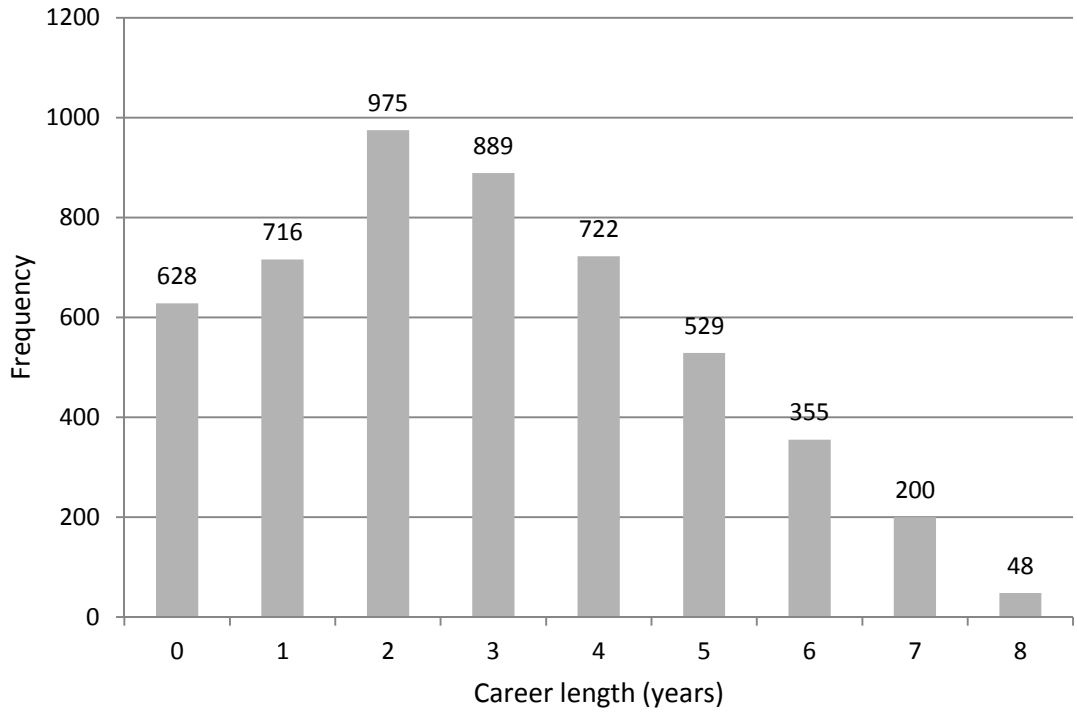

Figure 6-1. Frequency histogram of the career lengths (date of first race to date of retirement) of horses that retired from racing in Hong Kong between September 1996 and December 2010.

Genetic and phenotypic correlations of CL with the health traits studied previously in the HKJC data are shown in Table 6-1. CL was found to be significantly genetically correlated with both OA and tendon injury (p-value <0.001, LRT between unconstrained model and model constrained to r_g =0.00001). Phenotypic correlations were found between CL and EIPH/epistaxis, ligament injury and OA.

Table 6-1. Genetic and phenotypic correlations of CL with health problems in HKJC data.

	r_g* (CI), p-value	$r_p^{\dagger}$ (CI)
EIPH/epistaxis	0.1121 (-0.1339 – 0.3581), 0.38	0.0709 (0.0425 – 0.0993)
Fracture	0.1982 (-0.1575 – 0.5539), 0.30	0.0035 (-0.0239 – 0.0309)
Ligament injury	0.0371 (-0.3567 – 0.4309), 0.89	0.0347 (0.0069 – 0.0625)
OA	0.4371 (0.2090 – 0.6652), 0.00	0.0945 (0.0663 – 0.1227)
Tendon injury	0.4743 (0.2164 – 0.7322), 0.00	0.0145 (-0.0139 – 0.0429)
Suspensory ligament injury	-0.1052 (-0.5640 – 0.3536), 0.67	0.0248 (-0.0030 – 0.0526)

*Genetic correlation;†phenotypic correlation. CI indicates confidence intervals (± 1.96 x standard error).

Figure 6-2 shows the mean EBV of CL per year of birth. Regression of EBVs on year of birth was equivalent to a decrease of -0.0010 ± 0.0004 grades per annum, indicating a small but significant worsening trend.

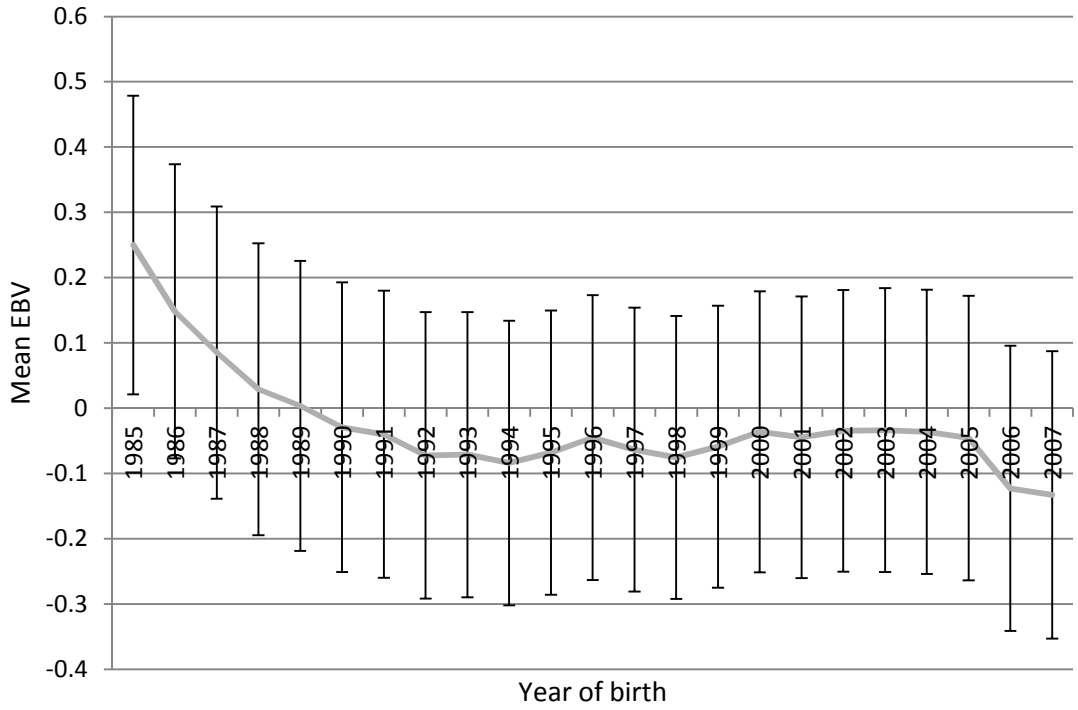

Figure 6-2. Mean EBV for CL by year of birth cohort. Error bars indicate mean EBV ± mean standard error per year of birth.

6.2.1.1.2 Discussion

In Hong Kong, based on our definition of CL, a horse's 'career' begins when it starts its first race in Hong Kong. No horses are bred in Hong Kong, therefore all racehorses originate from different countries, and may or may not race before being imported. Typically, flat racing horses begin their careers at 2 years of age, but only 3% of HK imports (156/5062) over the study period were 2 years old, thus many may have raced previously (data not shown). Also, no account was taken of training that occurred before the first race, as details of this were not available, thus the CL variable does not necessarily include all time spent at risk of exercise-related problems. Retirement was taken to be the end of the career of HK Thoroughbreds. Horses may be retired at any time, based on health, performance, or management grounds.

CL was found to be a significantly heritable trait, with a moderate heritability of 0.22, indicating a significant genetic component. Horses with longer careers were found to be more likely to suffer some musculoskeletal conditions, which is not surprising given the age-related nature of many of these traits, and increased time at risk. Breeding for longer career lengths could be possible, although greater emphasis on the avoidance and treatment of OA and tendon injuries would need to be encouraged.

The number of starts in Hong Kong over the career was not included in the final model of CL. Horses that run many races may be those of lower ability, if perceived 'good' horses are saved for infrequent but more prestigious races, or the alternative may be true, and 'good' horses may be entered into as many races as possible to maximize the chances of earning prize money. The relationship of CL with number of starts will be discussed further below.

6.2.1.2 BHA career length heritability analyses

6.2.1.2.1 Results

In the BHA PH dataset, 'career length' was calculated by subtracting the first from the last race date, plus 1, and these data were logarithmically transformed to account for their non-Normal distribution (CL). Career length was distributed as shown in Figure 6-3. The overall mean career length for all horses was 1.53 years, and 59.8% of horses started racing as 2 or 3 year olds (data not shown). Therefore, to ensure that information on career length included a majority of horses that had completed their careers, the PH dataset was reduced to include only horses that were born in 2005 or earlier (n=70387).

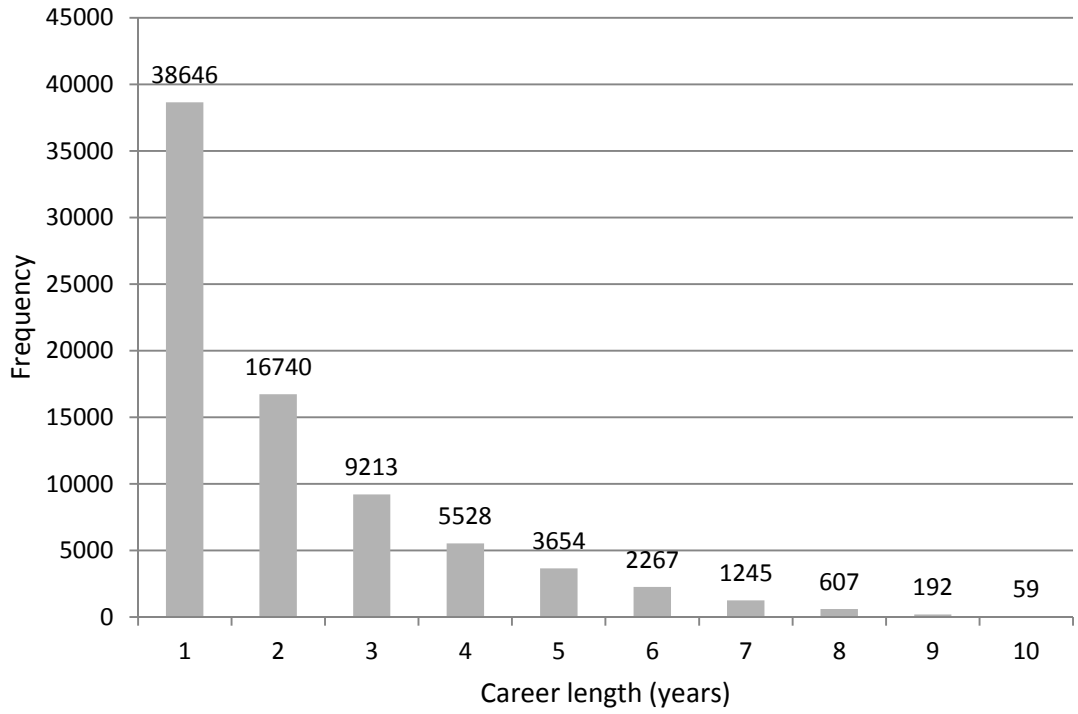

Figure 6-3. Distribution of career lengths (date of first race to date of last race) of horses in the BHA PH dataset.

Fixed effects found to be significantly associated with CL were the intensity of racing (number of starts divided by the career length in days), racing profile (flat, jump or mixed), sex, year of birth, and an interaction between racing intensity and year of birth (*p*-values all <0.001). The inclusion of animal as a random effect significantly improved the fit of the final model (*p*-value <0.001). The heritability of CL using a linear animal model was estimated to be 0.1630 (s.e. 0.0091, 95% confidence intervals 0.1452 to 0.1808).

Genetic and phenotypic correlations between CL, distal limb fracture, epistaxis and SDFT injury are shown in Table 6-2. CL was significantly negatively genetically correlated with SDFT injury (*p*-value <0.001, LRT between unconstrained model and model constrained to r_g =0.00001). CL was positively phenotypically correlated with all three conditions.

Table 6-2. Genetic and phenotypic correlations between CL and important conditions in the BHA PH reduced dataset.

	r_g* (CI), *p*-value	r_p† (CI)
Distal limb fracture	0.2017 (0.0177 – 0.3857), 0.06	0.0117 (0.0043 – 0.0191)
Epistaxis	0.0759 (-0.0627 – 0.2145), 0.33	0.0944 (0.0870 – 0.1018)
SDFT injury	-0.2998 (-0.4005 – -0.1991), 0.00	0.0628 (0.0552 – 0.0704)

*Genetic correlation;†phenotypic correlation. CI indicates confidence intervals (± 1.96 x standard error).

Figure 6-4 shows mean EBVs per year of birth cohort between 1983 and 2005 for CL. Regression of EBVs on year of birth was equivalent to an increase of 0.0052 ± 0.0001 grades per annum, indicating a small improving trend, but this was not found to be statistically significant.

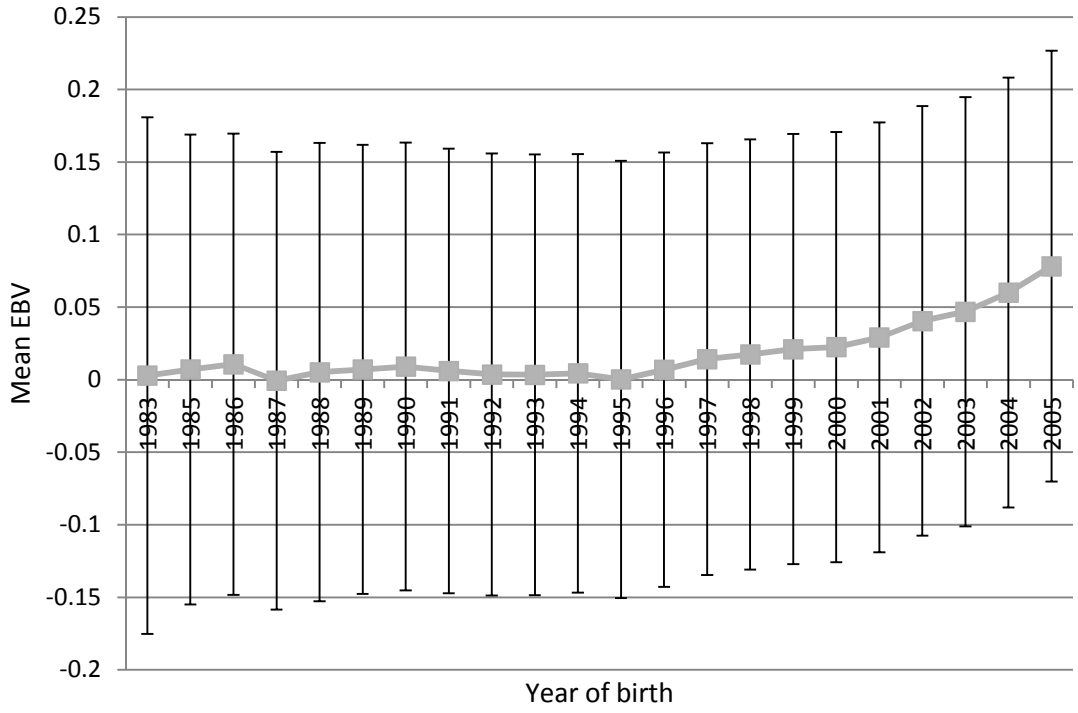

Figure 6-4. Mean EBV for CL by year of birth cohort. Error bars are mean EBV ± mean standard error per year of birth.

6.2.1.2.2 Discussion

CL in the BHA PH reduced dataset was associated with intensity of racing. As the number of career starts per unit time increased, the CL reduced. Racing intensity was a variable generated using career length as the ratio denominator, thus it should be interpreted with caution (see Chapter 2.1.2). CL was also associated with racing profile, where jumping or mixed racing careers were generally longer than flat racing careers. Females were found to have shorter careers than males (mean 1.1 years versus mean 1.9 years) (data not shown).

CL in this BHA dataset was found to be moderately heritable (0.16). This heritability is of a magnitude that selective breeding could be implemented to alter career length in this

population, and due to a significant genetic correlation, it is likely that SDFT injury would become less common if racing careers became longer.

No trend in EBVs per year of birth could be identified as standard errors spanned zero, however, mean EBVs appeared to increase towards the latter years. This trend was not significant.

6.2.2 Number of career starts

6.2.2.1 HKJC number of starts heritability analyses

6.2.2.1.1 Results

The number of starts made (in Hong Kong) by each horse before its retirement was available in the HKJC data, and was distributed as shown in Figure 6-6. These data were not Normally distributed, and 628 horses had no starts in Hong Kong before retirement, therefore the variable was grouped as follows: group 1 = no starts (n=628), group 2 = 1 to 14 starts (n=1540), group 3 = 15 to 29 starts (n=1277), and group 4 = more than 30 starts (n=1617). This variable will be referred to as NS.

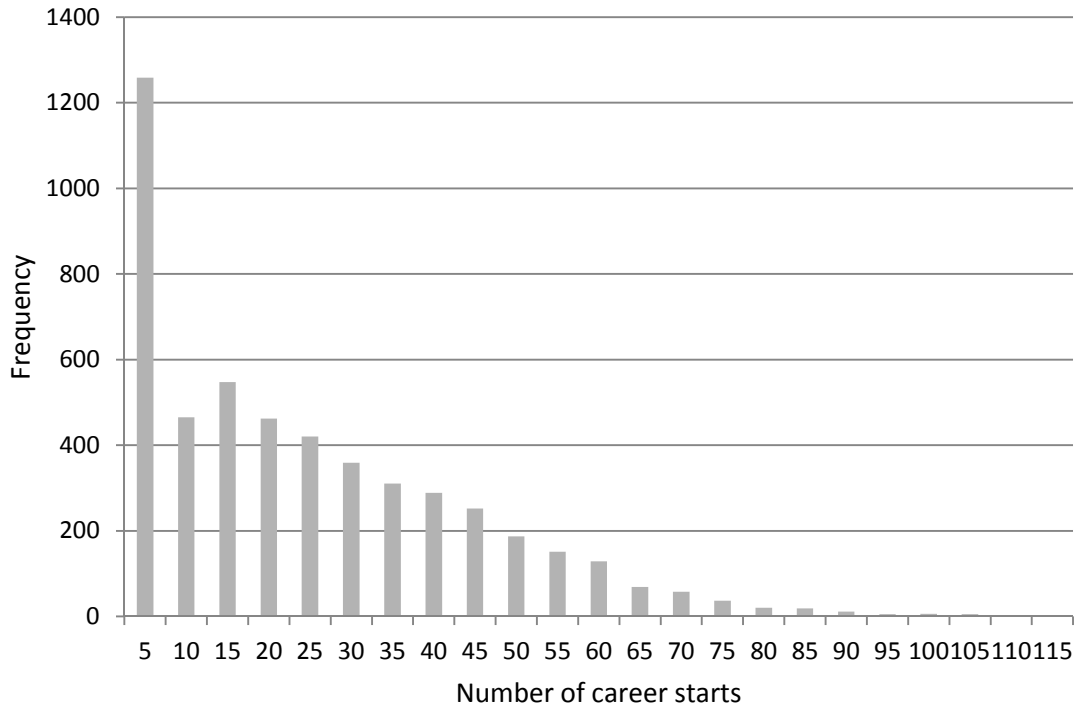

Figure 6-5. Frequency histogram of the number of starts per horse over the racing career in Hong Kong.

NS was modelled as described in Chapter 2 using a linear animal model. The only fixed effect found to be significant in the final NS model was the length of career (years). The heritability of NS was estimated to be 0.2556 (s.e. 0.0432, 95% confidence intervals 0.1709 to 0.3403).

Table 6-3 contains genetic and phenotypic correlations between NS and important medical conditions. No significant genetic correlations were found (*p*-values >0.05). NS was positively phenotypically correlated with EIPH/epistaxis, ligament injury and suspensory ligament injury, and was negatively phenotypically correlated with tendon injury.

Table 6-3. Genetic and phenotypic correlations between NS and selected medical conditions in the HKJC dataset.

	r_g* (CI), *p*-value	$r_p^{\dagger}$ (CI)
EIPH/epistaxis	-0.3525 (-0.7898 - 0.0848), 0.11	0.1049 (0.0773 – 0.1325)
Fracture	0.6169 (0.0193 – 1.2145), 0.05	0.0255 (-0.0021 – 0.0531)
Ligament injury	-0.0722 (-0.7606 – 0.6162), 0.85	0.0506 (0.0230 – 0.0782)
OA	0.3222 (-0.1651 – 0.8095), 0.20	0.1113 (0.0839 – 0.1387)
Tendon injury	-0.1118 (-0.6275 – 0.4039), 0.70	-0.0493 (-0.0769 - -0.0217)
Suspensory ligament injury	-0.1569 (-0.9615 – 0.6477), 0.72	0.0564 (0.0288 – 0.0840)

*Genetic correlation;†phenotypic correlation. CI indicates 95% confidence intervals (± 1.96 x standard error).

Figure 6-6 shows mean EBVs per year of birth for the number of career starts in Hong Kong. Regression of EBVs on year of birth was equivalent to an increase of 0.0124 ± 0.0004 grades per annum, indicating a small but significant increasing trend.

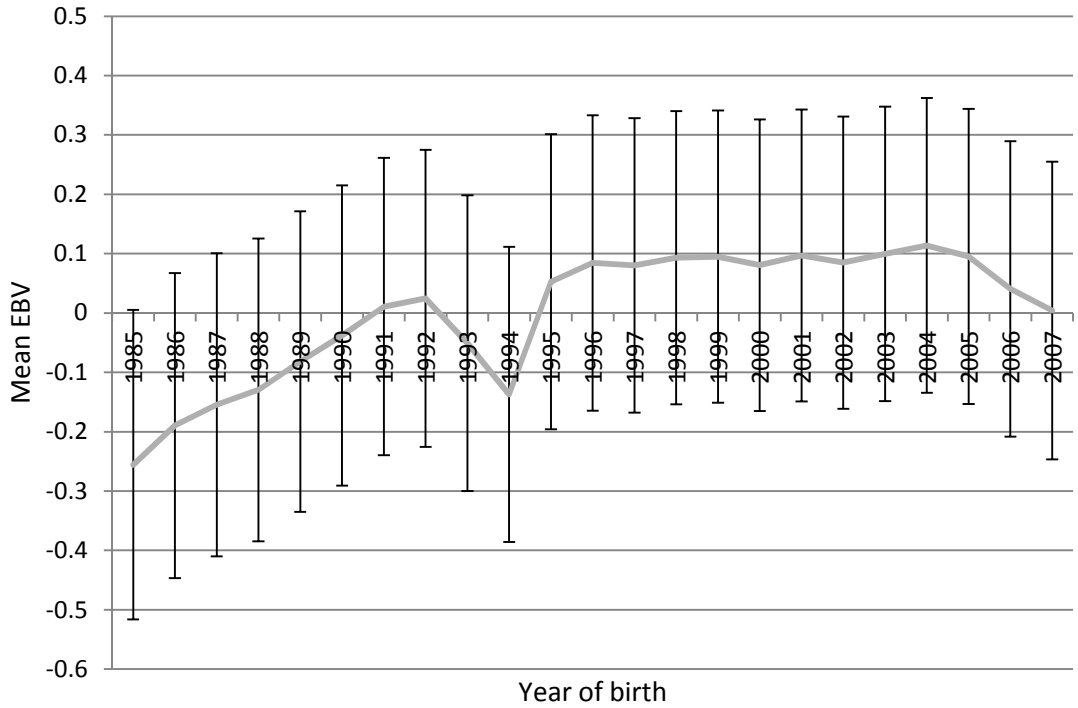

Figure 6-6. Mean EBV for NS in HKJC data by year of birth. Error bars indicate ± 1.96 x standard error.

6.2.2.1.2 Discussion

Many horses (n=628) in the HKJC data were not recorded as having started any races over their Hong Kong racing career. Because of this, it was thought best to group the starts data into a number of levels, one of which contained only horses that had no starts (Tozaki et al., 2012). This grouped variable had a moderate significant heritability (0.26), and was not correlated with any of the important health traits studied. The number of starts a horse is capable of making could potentially be increased using selective breeding strategies based on EBVs of NS, without significantly increasing the genetic risk of any of the medical conditions studied here.

The number of starts a horse completes over its career depends on many factors, and one found to be significant in the final model of NS here, was the length of career. A horse that fails to start many (or any) races is likely to be retired early, and therefore have a short career, and will also, therefore, be less likely to be subjected to many OVEs as it spends little time at risk.

6.2.2.2 BHA PH number of starts heritability analyses

6.2.2.2.1 Results

The number of starts for each horse was calculated from the BHA REP dataset, by summing the number of records for each horse. This variable was distributed as shown in Figure 6-7. Due to this non-Normal distribution, the number of starts was log transformed, before model building commenced as per Chapter 2, again using only horses that were born before 2006 to ensure most horses had completed their careers. This log-transformed version of the variable will be referred to as NS.

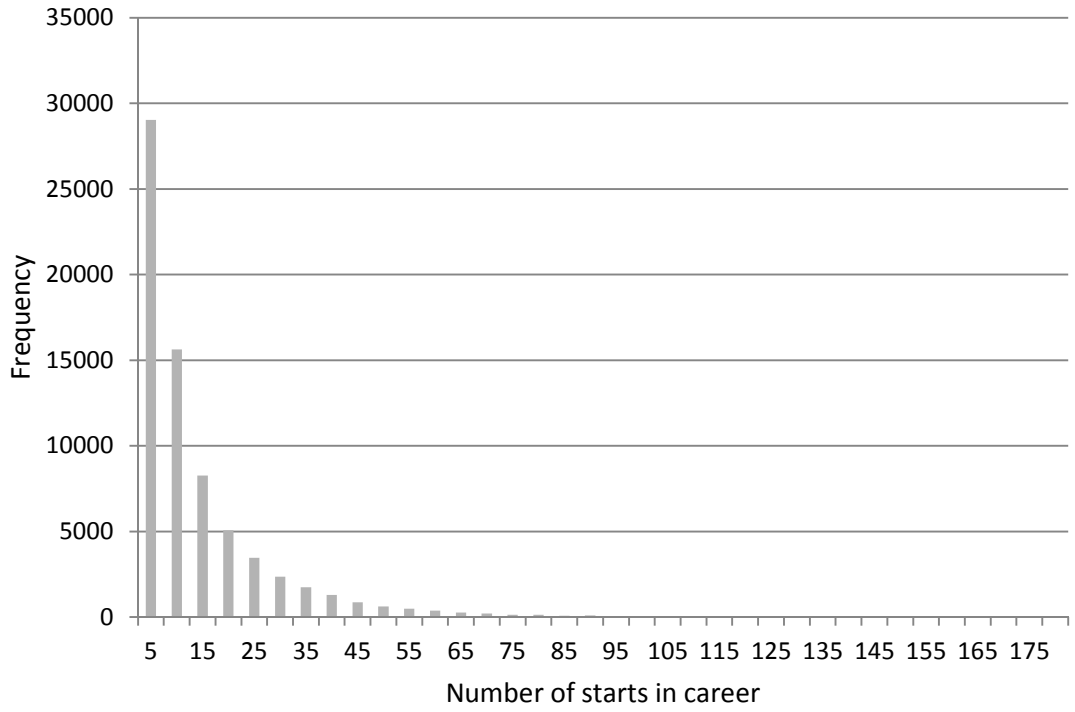

Figure 6-7. Frequency histogram of the number of starts per horse over the racing career.

Addition of animal as a genetic random variable to the final model of NS significantly improved model fit (*p*-value <0.001). Fixed effects found to be significantly associated with NS were intensity of racing (number of starts divided by the career length in days), racing profile (flat, jump or mixed), sex, year of birth, and first order interactions between intensity and sex, intensity and year of birth, racing profile and year of birth, and racing profile and sex (all *p*-values <0.05). The heritability of NS was estimated to be 0.2933 (s.e. 0.0101, 95% confidence intervals 0.2735 to 0.3131).

Genetic and phenotypic correlations between NS and important health conditions in the BHA data are shown in Table 6-4. The only significant (*p*-value <0.001) genetic correlation was found between NS and SDFT injury, and was negative in sign.

Table 6-4. Genetic and phenotypic correlations between NS and important conditions in the BHA PH reduced dataset.

	r_g* (CI), *p*-value	r_p† (CI)
Distal limb fracture	0.0301 (-0.1324 – 0.1926), 0.75	0.0172 (0.0098 – 0.0246)
Epistaxis	-0.0090 (-0.1297 – 0.1117), 0.89	0.1130 (0.1056 – 0.1204)
SDFT injury	-0.3828 (-0.4702 - -0.2954), 0.00	0.0487 (0.0411 – 0.0563)

*Genetic correlation;†phenotypic correlation. CI indicates confidence intervals (± 1.96 x standard error).

Figure 6-8 shows mean EBVs for NS per year of birth cohort. Regression of EBVs on year of birth was equivalent to an increase of 0.0089 ± 0.0001 grades per annum, indicating a small but significant increasing trend.

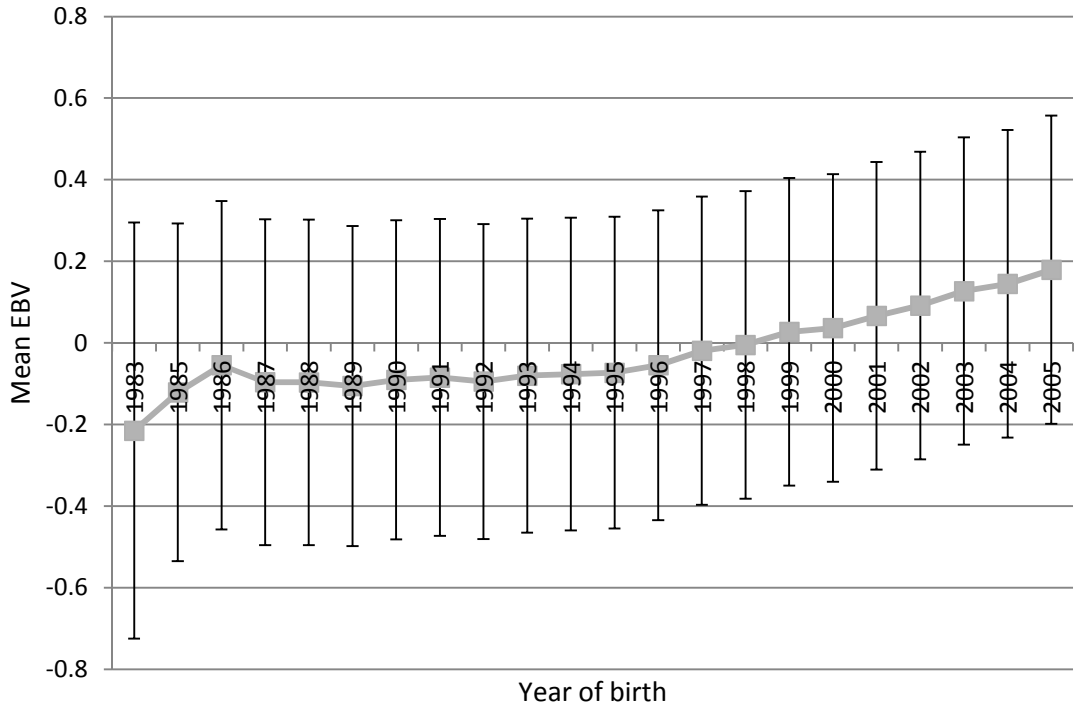

Figure 6-8. Mean EBV for NS in BHA PH dataset restricted to horses born before 2006. Error bars indicate ± 1.96 x standard error.

6.2.2.2.2 Discussion

The number of starts in a career, modelled as a log-transformed variable, was found to be moderately heritable at 0.29. Racing intensity was calculated using the number of starts as the ratio numerator, therefore interpretation of this significant fixed effect should be approached with caution. Flat racing horses ran an average of 12 races over their careers, and jump racing horses ran 8 races on average (data not shown). Male horses ran more races than females (mean of 14 compared to 9).

If selective breeding for increased number of career starts were to be performed, a concurrent reduction in the genetic risk of SDFT injury may be affected, due to the negative genetic correlation found between these traits.

6.3 Performance criteria

6.3.1 Career earnings

6.3.1.1 HKJC earnings heritability analyses

6.3.3.1.1 Results

Career earnings for each horse in the HKJC dataset were available in Hong Kong Dollars (HKD). A large number of horses earned nothing, therefore the variable was grouped as per work by Tozaki et al. (2012). Groupings were as follows: group 1 = no earnings (n=1043), group 2 = 1 to <1.5 million HKD (n=2267), group 3 = 1.5 million to <3 million HKD (n=1075), group 4 = 3 million HKD and over (n=677). This grouped variable will be referred to as WG. Model building was performed as per Chapter 2.

Seven fixed effects were found to be significantly associated with WG (*p*-value <0.05). Fixed effects included age at retirement (years, with 10 and 11 grouped together), number of career starts, length of career (years), intensity of racing (number of starts divided by career length in days), year of retirement, trainer at the time of retirement, and neuter status (neutered or not neutered). A higher WG was found to be associated with older ages at the time of retirement, having started more races over the career and retiring in later years of the study. WG was found to be negatively associated with the career length and racing intensity, and being female.

Inclusion of animal as a random effect significantly improved the fit of the final model of WG (*p*-value <0.001). Heritability of WG was estimated to be 0.1232 (s.e. 0.0308, 95% confidence intervals 0.0628 to 0.1836).

Genetic and phenotypic correlations between WG and important health traits are given in Table 6-5. WG was found to be significantly negatively genetically correlated with EIPH/epistaxis.

Table 6-5. Genetic and phenotypic correlations between WG and selected medical conditions in the HKJC dataset.

	r_g* (CI), *p*-value	$r_p^{\dagger}$ (CI)
EIPH/epistaxis	-0.5007 (-0.8551 - -0.1463), 0.01	0.0526 (0.0246 – 0.0806)
Fracture	-0.0630 (-0.5620 – 0.4360), 0.81	-0.0140 (-0.0416 – 0.0136)
Ligament injury	-0.2229 (-0.7619 – 0.1361), 0.44	0.0655 (0.0379 – 0.0931)
OA	-0.2875 (-0.6215 – 0.0465), 0.09	0.1204 (0.0926 – 0.1482)
Tendon injury	-0.3139 (-0.6975 – 0.0697), 0.15	-0.1200 (-0.1474 - -0.0926)
Suspensory ligament injury	-0.1513 (-0.7805 – 0.4779), 0.66	0.0742 (0.0466 – 0.1018)

*Genetic correlation;†phenotypic correlation. CI indicates 95% confidence intervals (± 1.96 x standard error).

Figure 6-9 shows mean EBVs per year of birth cohort for WG. Regression of EBVs on year of birth was equivalent to an increase of 0.0013 ± 0.0003 grades per annum, indicating a small but significant improving trend.

Figure 6-9. Mean EBV for WG per year of birth cohort in HKJC data. Error bars indicate ± 1.96 x standard error.

6.3.3.1.2 Discussion

A horse can only accrue earnings if it is entered in races, and is placed in at least one of those races. The magnitude of prizes for each race are determined by economic factors, and by the quality of the race. To some degree, therefore, environmental effects will determine how much a horse is likely to earn. Despite this, WG was significantly heritable in this population, with a moderate heritability of 0.12. Due to the negative genetic correlation between EIPH/epistaxis and WG, if WG EBVs were to be used as selection criteria to selectively breed for increased WG, a reduction in EIPH/epistaxis may be seen concurrently. This finding supports the theory that EIPH/epistaxis is associated with poor performance, but is the first time this link has been shown to be at least partly due to common genetic factors.

Horses that were older at the time of retirement would be more likely to have run more races, and thus had more chances to accrue prize money over their careers. Having more career starts, and retiring at older ages were associated with increased career earnings in the present study. The relationship between the length of career and WG was found to be negative, however, therefore shorter careers were associated with higher winnings in this population. It is hypothesized that this relationship occurs because the most profitable racehorses are sent to stud early to obtain maximum returns on breeding, or that it was relatively common practice to import horses into Hong Kong for a short time to enter high stakes races only, leading to short 'careers' in Hong Kong but potentially greater earnings.

The identity of the trainer at the time of retirement was significantly associated with WG. Although trainers undoubtedly have an important effect upon the success of the horses under their care, this relationship should be interpreted with caution, because each horse could have multiple trainers over its career, each for a variable amount of time. The effects that a trainer has on a horse's ability to win are likely to be cumulative to some degree, thus the current trainer is not necessarily the one that had most influence over a horse's performance phenotype at any one time. The association between neuter status and WG should also be interpreted cautiously, as 99% (5028/5062) of horses were male, and of those 94% (4740/5028) of horses were neutered.

Later years of retirement were associated with higher earnings. It is thought that this is probably partly an economic effect of the availability of prize money, as well as being due

to a marginal improvement in horse performances as recorded by an increasing trend in WG over the study period.

6.3.3.2 BHA earnings heritability analyses

6.3.3.2.1 Results

Lifetime earnings (in British pounds) for each horse were available in the BHA data. Horses born in 2006 or later were excluded to ensure that those remaining were likely to have concluded their careers. Horses were grouped according to their lifetime earnings as per the work by Tozaki et al. (2012). Group 1 had no earnings (n=33417), group 2 earned between £1 and £4999 (n=11194), group 3 earned between £5000 and £49,999 (n=21794), and group 4 earned £50,000 or more (n=3982). This variable will be referred to as WG. Model building was carried out as per Chapter 2.

Fixed effects found to be significantly (*p*-value <0.001) associated with WG were the number of career starts, sex, year of birth and racing profile (flat, jump, or mixed racing). Inclusion of animal as a random effect significantly improved the fit of the final model (*p*-value <0.001). Heritability of WG was estimated to be 0.3564 (s.e. 0.0108, 95% confidence intervals 0.3352 to 0.3776).

Genetic and phenotypic correlations between WG, distal limb fracture, epistaxis and SDFT injury are shown in Table 6-6. WG was positively genetically correlated with distal limb fracture and SDFT injury (*p*-value <0.05, LRT between unconstrained model and model constrained to r_g =0.00001).

Table 6-6. Genetic and phenotypic correlations between WG and important conditions in the BHA PH reduced dataset.

	r_g* (CI), *p*-value	r_p† (CI)
Distal limb fracture	0.1930 (0.0346 – 0.3514), 0.02	0.0213 (0.0139 – 0.0287)
Epistaxis	0.1146 (-0.0005 – 0.2297), 0.06	0.0651 (0.0575 – 0.0727)
SDFT injury	-0.1048 (-0.1861 - -0.0235), 0.01	0.0219 (0.0141 – 0.0297)

*Genetic correlation;†phenotypic correlation. CI indicates confidence intervals (± 1.96 x standard error)

Figure 6-10 shows the mean EBV for WG in the BHA data per year of birth. Regression of EBVs on year of birth was equivalent to an increase of 0.0209 ± 0.0003 grades per annum, indicating a small but significant improving trend.

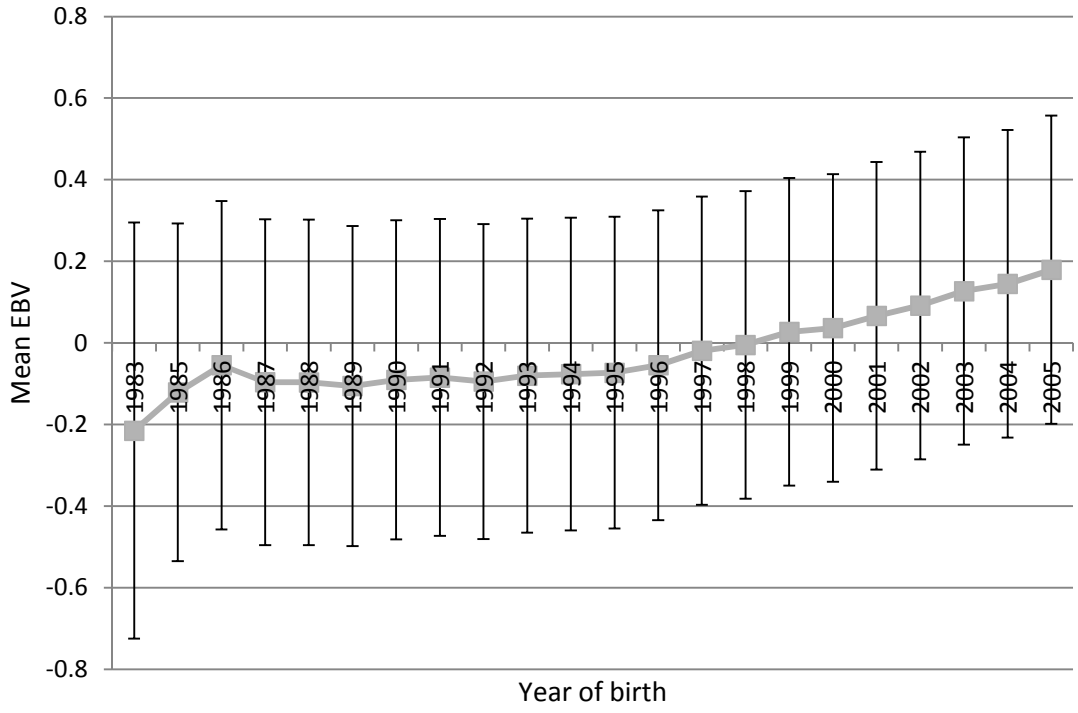

Figure 6-10. Mean EBV for WG per year of birth cohort in BHA PH data. Error bars indicate ± 1.96 x standard error.

6.3.3.2.2 Discussion

The heritability of WG in the reduced BHA PH dataset was high at 0.36. The amount of available prize money in any one year is determined by many economic and human factors; however, the career earnings of each horse are strongly genetically determined.

WG was positively genetically correlated with distal limb fracture, suggesting that if career earnings were to be maximized through selective breeding based on EBVs of WG, the risk of these fractures may increase. The magnitude of this correlation was moderate at 0.19 with wide standard errors. The heritability of distal limb fracture was found to be low in this study (see Chapter 3), thus it is hoped that, should WG be selectively bred for, environmental manipulations could avoid an actualized increase in distal limb fracture incidence. Such breeding strategies may also benefit horse welfare by reducing SDFT injury risk concurrently, although the size of the genetic correlation between SDFT injury and WG was small at -0.10.

6.3.4 Durability and earnings compared

6.3.4.1 HKJC data

Table 6-7 shows Pearson correlation coefficients between the three measures of durability and performance used in the HKJC data. All of the variables were positively correlated, with the greatest correlation between CL and NS.

Table 6-7. Pearson correlation coefficients between CL (career length, see 6.2.1.1), NS (number of career starts, see 6.2.2.1) and WG (earnings group, see 6.2.3.1). Figures in parentheses are *p*-values.

	CL	**NS**	**WG**
CL			
NS	0.81 (<0.001)		
WG	0.42 (<0.001)	0.60 (<0.001)	

Table 6-8 shows the genetic and phenotypic correlations between the same three durability or performance variables, with their heritability estimates as shown earlier. None of the genetic correlations were significant at *p*-value of <0.05. CL and NS, and WG and NS were positively phenotypically correlated with each other.

Table 6-8. Heritability estimates (on diagonal), genetic correlations (below diagonal, with their *p*-values italicised following the estimate) and phenotypic correlations (above diagonal) between CL (career length, see 6.2.1.1), NS (number of career starts, see 6.2.2.1) and WG (earnings group, see 6.2.3.1). Confidence intervals are ± 1.96 x standard error.

	CL	**NS**	**WG**
CL	0.2167 (0.1383 – 0.2951)	0.0562 (0.0284 – 0.0840)	0.0096 (-0.0186 – 0.0378)
NS	0.3351, *0.16* (-0.0947 – 0.7649)	0.2556 (0.1709 – 0.3403)	0.8127 (0.8033 – 0.8221)
WG	-0.1413, *0.45* (-0.4782 – 0.1956)	0.6734, *0.10* (0.3782 – 0.9686)	0.1232 (0.0628 – 0.1836)

6.2.4.2 BHA data

Table 6-9 contains Pearson correlation coefficients between CL, NS and WG in the BHA PH reduced dataset. All correlations were significantly greater than zero, and positive in sign. CL and NS were highly correlated.

Table 6-9. Pearson correlation coefficients between CL (career length, see 6.2.1.2), NS (number of career starts, see 6.2.2.2) and WG (earnings group, see 6.2.3.2). Figures in parentheses are *p*-values.

	CL	NS	WG
CL			
NS	0.85 (<0.001)		
WG	0.34 (<0.001)	0.46 (<0.001)	

Table 6-10 shows genetic and phenotypic correlations between CL, NS and WG with their heritability estimates as calculated previously. All variables were genetically and phenotypically correlated with each other.

Table 6-10. Heritability estimates (on diagonal), genetic correlations (below diagonal, with their *p*-values italicised following the estimate) and phenotypic correlations (above diagonal) between CL (career length, see 6.2.1.2), NS (number of career starts, see 6.2.2.2) and WG (earnings group, see 6.2.3.2). Confidence intervals are ± 1.96 x standard error.

	CL	NS	WG
CL	0.1630 (0.1452 – 0.1808)	0.8480 (0.8458 – 0.8502)	0.3359 (0.3286 – 0.3432)
NS	0.2485, *<0.001* (0.1476 – 0.3494)	0.2933 (0.2735 – 0.3131)	0.4627 (0.4562 – 0.4692)
WG	0.3517, *<0.001* (0.2931 – 0.4103)	0.3994, *<0.001* (0.3541 – 0.4447)	0.3564 (0.3352 – 0.3776)

6.4 Discussion

Pearson correlation coefficients between CL, NS and WG showed that all three of these variables have a tendency to increase linearly with each other, although this relationship was less clear between CL and WG. Bivariate linear animal models were constructed to assess whether these correlations could be due to genetic factors. No significant genetic correlations between CL, NS or WG were found in the HKJC data, but all of these variables were significantly (positively) genetically correlated with each other in the BHA data. The size of these correlations were moderate, therefore it is likely that a proportion of the genes that are common to each variable are shared. In Hong Kong, the previous racing careers of imported horses were not included in these analyses. In the BHA data, it was not certain when a horse had finished its racing career, therefore only horses born

before 2006 were included in the analyses. In both datasets, the time spent training before the first race was started was not included.

The length of 'career' was investigated here to assess whether it was a variable associated with health or with performance, and if so, whether it was heritable. CL was found to be heritable in both study populations (0.22 HKJC, 0.16 BHA), despite the somewhat different definitions of CL in these groups. Longer careers were associated with greater earnings and more starts, although this relationship did not appear to be due to genetic factors in the HKJC data. In the BHA data, CL was genetically correlated with both NS (0.25) and WG (0.35). Longer career lengths in Hong Kong were genetically correlated with OA and tendon injury, suggesting that breeding to increase CL would increase risk of these conditions. In the BHA data, longer career lengths were genetically correlated with reduced risk of SDFT injury. These findings suggest that the genes associated with CL may be different in BHA and HKJC populations. CL is not an appropriate target for selective breeding in Hong Kong, as the current study has shown that it is not correlated with performance-related traits (WG), and may increase the risk of OA and tendon injury. Conversely, in Britain, breeding to increase CL may lead to a reduction in SDFT injury, and greater career earnings. The heritability of CL in BHA data was, however, lower than the heritability of WG or NS, thus these may be more appropriate breeding targets. In both datasets, the activity of horses during their 'career' was not recorded, and it must be noted that those horses undergoing active race training may be placed at higher risk of adverse health events than those horses that are not being actively trained for variable periods of time.

Horses capable of running more races would be of benefit to horse owners and trainers, where the chances of recouping a proportion of the costs of housing and training would be increased. It is not clear whether the best horses are reserved for fewer more prestigious races, or whether horses of high ability are entered in as many races as possible. Decisions such as these rely on human decision making and individual economic circumstances. It can be supposed, however, that only healthy horses would be capable of running many races, thus increasing the number of starts each horse can make is of veterinary interest. NS was significantly heritable in both datasets used here (0.26 HKJC, 0.29 BHA), suggesting a large genetic component to this variable. NS does not appear to increase the genetic risk of any of the health conditions studied, and may in fact reduce the risk of SDFT injury as NS increases. The genetic factors underlying NS also act to increase WG

and CL in the BHA data. Horse ability, as related to lifetime earnings, is not associated with the number of starts in HK, but in the BHA data it appears that horses of greater ability run more races. These findings combined suggest that selective breeding for increased NS in either HK or BHA may be possible, and of benefit to the horse population and stakeholders.

Prize money is generally only available for horses that finish first to third in races. To win or be placed in a race depends on the horse's own performance, but also the performance of its competitors, and how the environment interacts with horse performance. The size of available cash prizes depends on the race, and on local economic circumstances. Some horses may not have the chance to earn large amounts if they are only entered in lower-class races, or are consistently raced with horses of greater ability. All of these factors could be supposed to blur the relationship between career earnings and horse ability. Despite this, WG was significantly heritable in HKJC data (0.12), and was highly heritable in BHA data (0.36). This latter estimate exceeds those of Tozaki et al. (2012), where lifetime earnings were found to have a heritability estimate of 0.25. In the HKJC data, longer career lengths were associated with lower career earnings, suggesting that horses with more ability are retired earlier, perhaps to be exported to stud or to racing yards elsewhere. There was no association found between career length and career earnings in the BHA data. Breeding for increased earnings is of obvious interest to owners, but should only be of interest to veterinarians if horse health is also improved. In the HKJC dataset, WG was significantly negatively genetically correlated with EIPH/epistaxis (-0.50), and in the BHA dataset WG was significantly negatively genetically correlated with SDFT injury (-0.10). Unfortunately, a positive genetic correlation between WG and distal limb fracture in the BHA data was found, suggesting that increasing WG would lead to increased risk of this condition. However, the heritability of distal limb fracture was found to be low in this population (see Chapter 3), so if selective breeding for increased WG were to be conducted, appropriate environmental modifications could be made to reduce individual fracture risk. Using the BHA data, breeding for increased WG would be more successful than breeding for either CL or NS, due to its high estimated heritability; however, all of these factors would be expected to increase thanks to positive genetic correlations between them.

Only a select few health traits were considered in these analyses, although they were deemed to be important to each population. There may be more positive correlations

between CL, NS or WG with adverse health traits that were not investigated in this study. For this reason, further work would need to be done to thoroughly investigate the possible outcomes of breeding for these performance and durability traits in the UK, and in those countries supplying racehorses to Hong Kong, in the future.

An alternative approach to selective breeding for performance or durability based on EBVs could be to offer commercial genotyping of panels of important associated SNPs, to inform future matings or evaluate young stock. Such commercial tests exist in the public domain currently, for assessing the most appropriate race distances for individual horses, and to attempt to identify horses capable of elite performances (www.equinome.com). Due to the polygenic nature of these traits, assessment of a panel of selected SNPs can only ever provide an estimate of the genetic value of an individual, and client education is essential to avoid misunderstandings surrounding the probability-based nature of test results.

7. CHAPTER VII. General discussion

Heritability analyses are an important component of disease research. They shed light on the sources of variation in a trait within populations, and allow comparison between populations and traits in terms of the relative importance of genetic influences. Following such analyses, research efforts can be focused on the most fruitful courses of action to manipulate risk, and studies at the molecular level can help investigate causative pathways and generate commercial applications for animal breeding. The generation of estimated breeding values as part of heritability analyses can be used to rank animals on relative genetic merit, and are a vital part of animal breeding for improved performance, production, or reduced disease risk (Hoque et al., 2007; Thomson et al., 2010; Webbon, 2012). These values can also be used to investigate the change in population risk of disease over time, or appreciate the effects of selective breeding interventions. Unlike food animals such as cattle and pigs, Thoroughbred horses in the UK have yet to be subjected to extensive selective breeding measures using genetic indices, but certain breeds in other countries are being 'improved' using EBVs for performance indicators (Bugislaus et al., 2006; Schroderus and Ojala, 2010; Suontama et al., 2012). Given the formidable size and importance of the Thoroughbred horseracing industry, genetic research and selective breeding are not only feasible, but could provide real benefits to horse health and performance, and may help safeguard the breed, the reputation of the veterinary profession and the reputation of the racing industry. These studies were conducted to investigate a number of diseases that are prevalent in racehorses, in the hope that the results would give clear indications of how best to direct efforts to reduce future risk. Additionally, a number of performance and longevity traits were subjected to heritability analyses, to assess whether they might be appropriate breeding targets, and avoid the need to select horses based on many individual disease traits concurrently. For most of the health traits studied, this report constitutes the first attempt to quantify the genetic contribution to risk in the Thoroughbred. For those conditions previously analysed elsewhere, the results reported here offer validation and interesting comparisons, and help shed light on whether the nature of genetic determination of conditions differs in different racing jurisdictions.

Two very different sources of data were used in this study. The HKJC supplied two sources of horse health information, gathered over an eighteen year period. These data reported diagnoses and investigations of health events, whether associated with exercise or rest periods. The reason(s) behind retirement was also provided for a large number of horses, which were not restricted to medical causes. The main benefit of using HKJC data for heritability analyses was that horses under the care of the HKJC could be expected to

experience a similar environment (feeding, housing, training etc.), during their time in Hong Kong. This similarity of environment may have reduced the environmental variance, compared to the supposedly more varied environment experienced by horses in the UK, for example. Such a reduction may have minimised phenotypic variance, leading to inflated heritability estimates. However, imported racehorses originated from a number of different countries, and arrived in Hong Kong at different stages of their career. Thus their previous environmental exposures would have been diverse, and the overall effect on heritability may have been modest. Some duplication of information occurred both within the OVE data, and between the OVE and retirement data. Because this could not be adequately rectified, analyses were conducted after reducing the data down to a single binary disease 'profile' per horse. Repeated measures were not available, and graduated scales of clinical significance were not reported in the original diagnoses. Binary data have been used extensively in disease studies in the past (Stock and Distl, 2006; Oki et al., 2007; van Grevenhof et al., 2009; Vazquez et al., 2009; Schurink et al., 2011; Pritchard et al., 2012). It has been frequently noted that animal diseases can often be expected to follow an unseen, underlying pattern of liability with a di- or polychotomous phenotypic outcome (Falconer and Mackay, 1996; Matos et al., 1997; Stock and Distl, 2006; Oki et al., 2008; Schurink et al., 2011). Using modelling techniques that take this liability into account often produces higher heritability estimates. When disease is recorded as either present or absent (binary), logistic regression can be employed to similar effect. Using the log odds of an outcome, based on prevalence, allows prediction of the effects of explanatory variables without values being restricted to 0 or 1.

HKJC data required the use of content analysis to produce an analysable numeric dataset for heritability analyses. This is a process commonly employed in data mining projects, where manual categorisation of text strings would be unfeasibly time consuming and prone to human error. The method was proven to be useful in the interrogation of equine health data from the same source by Lam et al. (2007). Careful population of the dictionaries used is essential, and without visual verification of every record, misclassification of a small proportion of records is possible. However, this risk was minimised by extensive use of the 'keyword in context' function of the content analysis software, and frequent verification of classifications. Human error in data input can reduce the sensitivity of content analysis, as spelling mistakes and alternative nomenclature can lead to classification errors or omissions. The content analysis procedures were conducted in an

iterative process in this study, until 100% of HKJC records had been classified, and no systematic errors in the dictionaries were evident.

Diagnoses of medical conditions included in OVE or retirement records were not accompanied by details of the methods of diagnosis. Whether visual diagnostic aids such as radiography or ultrasonography were involved is, therefore, unknown. The validity of diagnoses made using clinical examination only would be expected to be variable, depending on the condition present and the severity of clinical signs. A displaced mid-body cannon bone fracture, for example, would be straightforward to diagnose, but a mild suspensory desmitis or fracture to an individual carpal bone would be much more difficult. The sensitivity and specificity of HKJC diagnoses were assumed to be low, giving rise to artificially small heritability estimates, and in some cases, a lack of clarity about the significance of genetic variance (Bishop and Woolliams, 2010). Validation of a proportion of the diagnoses would have supported our use of these data for heritability analyses, but was not possible. Also, over the study period, there may have been changes in veterinary personnel, diagnostic capabilities and alterations in the prevalence of the conditions studied, none of which were incorporated here.

The second dataset used in this study was provided by the BHA, and contained information on every start on British racecourses between 2000 and 2010. No information on horse health away from the racetrack was available. Health information provided was in the form of binary 'profiles' for each horse following each start, detailing the presence or absence of a diagnosis of distal limb fracture, epistaxis, or SDFT injury at that race. Again, information on the severity of these conditions was not available, and most diagnoses would have been made without the use of technologies such as radiography, thus errors in diagnosis may have reduced sensitivity and specificity. However, repeated measures were available, as most horses completed a number of starts over the study period.

The method used to estimate heritability and genetic correlations for both datasets in this study was residual maximum likelihood (REML). This is a powerful method employed in animal breeding programmes, which is capable of estimating variance components despite very unbalanced data. An alternative approach to REML for heritability and genetic correlation estimation is the Markov-Chain Monte Carlo Bayesian (MCMC) approach using Gibbs sampling. This method has been employed in horse genetic studies in the

past, but REML remains the most popular method due to its flexibility and computational ease (Oki et al., 2008). Comparative studies of Bayesian and frequentist approaches to heritability analysis have suggested that both methods are susceptible to bias, particularly in cases of extreme prevalence, heritability or small sample size (Stock et al., 2007). Using univariate models of binary traits, as was mostly the case in this study, Gibbs sampling and REML produce similar estimates, and animals are ranked similarly on genetic merit. Gibbs sampling may outperform REML where multivariate models are needed, particularly where binary or categorical traits are involved. In situations where trait heritability is high, prevalence is low and the population is small, breeding values from Gibbs sampling are more reliable than those from REML (Stock et al., 2007). In most cases, the conditions studied here were of low to moderate heritability, prevalence was low or moderate, and population size was moderate or large, thus it can be assumed that REML was a suitable and reliable method of estimation.

Estimated breeding values derived from animal model analyses allow comparison of the relative genetic merit of individuals for the trait being studied. An EBV value describes how the potential offspring of an animal would compare to the mean of the population with respect to a given trait. The usefulness of EBVs is limited by their accuracy, which describes how close to the 'true' hierarchy of genetic merit the rankings of animal EBVs are, i.e. how close to the breeding value they are. EBVs of low accuracy approach the mean of the population, whereas the higher the accuracy, the more extreme (positive or negative) the EBV values may become, and the more confidence we may have that selection based on these EBVs could be efficient. In general, the EBVs produced in the analyses of this study were of low to moderate accuracy, and as such would not be of great use in producing rapid change in genetic merit of the populations studied.

Accounting for sources of similarity between individuals in a population is a key component in heritability analyses. Individual horses in a racing jurisdiction are not independent units, because they are related to some degree through their pedigree. REML allows the inclusion of this source of non-independence by calculating the expected proportion of genes shared by every pair of individuals in the pedigree. The tendency for cases or controls to be more common among more related individuals is then evaluated. Two ways of approaching this evaluation are to use sire models, or animal models. Sire models contain the identity of the sire as a random variable, thus half-sibling groups are identified and expected to share 25% of their genetic material. Dam relationships are not

included, sires are assumed to be unrelated to each other, and random mating is assumed. Sire models are widely implemented in animal breeding and genetics research, due to their computational ease and appropriateness in populations where there are few offspring per dam (Gianola, 1980; Matos et al., 1997; Sun et al., 2009; Stewart et al., 2012). Alternatively, animal models use all relationships in a pedigree to account for non-independence of individuals. Animal models can be considerably more computationally intensive than sire models, and convergence problems can be an issue. Whether sire or animal models produce larger heritability estimates varies depending on the population structure, and whether sires or dams are under greater selection pressure. Genetic correlations using sire or animal models tend to differ in magnitude but are similar in significance and direction (Stock and Distl, 2006).

Both sire and animal models can be constructed on the observed or logistic scales, to account for the binary nature of the data. Logistic models tend to produce larger heritability estimates, as they suffer less loss of precision and can capture more of the genetic variance of a trait. For categorical dependent variables, post-hoc conversion of heritability estimates on the observed scales can be made as per Dempster and Lerner (1950):

$$h_c^2 = h_o^2 \left(\frac{1-p}{i^2 p} \right)$$

where h_c^2 is heritability on the continuous liability scale, h_o^2 is heritability on the observed binary scale, p is the prevalence of the condition, and i is the mean liability of individuals with the condition in question at prevalence p, in units of standard deviation from the population mean (Dempster and Lerner, 1950). Heritability estimates on the continuous liability scale in this study, based on this conversion, occasionally exceeded 1, and often produced very large estimates, especially where prevalence was low and heritability on the observed scale was moderate or high (see Table 7-1). It was thought preferable to adjust for the binary data nature by using logistic regression, rather than this post-hoc conversion for the conditions studied here.

Table 7-1. Comparison of heritability estimates based on three methods.

Dataset	Data subset	Condition	Model form	Prevalence (%)	h^2_o*	h^2_l†	h^2_c‡
HKJC		Fracture	Sire	13	0.04	0.11	0.02
			Animal	13	0.03	0.05	0.02
		Suspensory ligament injury	Sire	10	0.06	0.17	0.17
			Animal	10	0.05	0.07	0.14
		Tendon injury	Sire	19	0.13	0.20	0.27
			Animal	19	0.14	0.09	0.30
		OA	Sire	10	0.03	0.15	0.09
			Animal	10	0.01	0.05	0.03
		EIPH/epistaxis	Sire	19	0.21	0.33	0.44
			Animal	19	0.19	0.13	0.40
BHA	All horses	Distal limb fracture	Sire	1	0.01	0.22	0.13
			Animal	1	0.01	0.16	0.14
		Epistaxis	Sire	2	0.02	0.26	0.18
			Animal	2	0.02	0.17	0.19
		SDFT injury	Sire	3	0.04	0.24	0.27
			Animal	3	0.18	0.14	1.24
BHA	Flat racing horses only	Distal limb fracture	Sire	1	0.01	0.20	0.12
			Animal	1	0.01	0.15	0.12
		Epistaxis	Sire	2	0.02	0.28	0.16
			Animal	2	0.02	0.18	0.16
		SDFT injury	Sire	1	0.05	0.52	0.54
			Animal	1	0.05	0.25	0.54
BHA	Jump racing horses only	Distal limb fracture	Sire	1	0.01	0.24	0.14
			Animal	1	0.01	0.16	0.15
		Epistaxis	Sire	3	0.03	0.26	0.19
			Animal	3	0.03	0.15	0.19
		SDFT injury	Sire	5	0.03	0.17	0.07
			Animal	5	0.03	0.11	0.07

*Observed linear scale;†Logistic scale; ‡ Linear scale estimate converted to underlying liability scale as per Dempster and Lerner (1950).

Heritability estimates derived from binary data on the observed scale are related to the prevalence of the condition, as can be seen in Table 7-1. Higher prevalence tends to lead to larger estimates, however estimates on the observed scale tend to be smaller than those from logistic models overall. More of the genetic variance can be captured when the logistic scale is used, as the predicted values are no longer restricted to being zero or one. This tends to lead to larger heritability estimates. In the present study, logistic estimates frequently exceeded linear estimates, as has been seen previously (Matos et al., 1997; Pieramati et al., 2003). Instances where logistic heritability estimates were smaller than those on the linear (observed) scale were restricted to animal models (tendon injury and EIPH/epistaxis analyses in HKJC). Table 7-2 contains examples of published literature using both animal and sire linear and logistic models in genetic analyses.

Table 7-2. Examples of the use of animal and sire models on linear and logistic scales in published literature.

Model form	Scale	Literature
Sire	Linear (observed)	Vazquez et al. (2009) Pritchard et al. (2012) Matos et al. (1997)
	Logistic	Vazquez et al. (2009)
Animal	Linear (observed)	van Grevenhof et al. (2009) Stock et al. (2007)* Lewis et al. (2010) Matos et al. (1997) Stock et al. (2006)* Pieramati et al. (2003)*
	Logistic	van Grevenhof et al. (2009) Suontama et al. (2012)

*Heritability estimates on the observed scale were transformed to the underlying liability scale as per Dempster and Lerner (1950).

Fractures sustained by racehorses are always of serious concern, and often fatal (McKee, 1995; Parkin, 2008). Whether they occur during racing, training or rest periods, their impact on horse welfare is considerable. Due to the large size and ungulate anatomy of the domestic horse, orthopaedic repair of many limb fractures is unfeasible, impossible, or undesirable due to the inevitable restriction of the animal's normal behaviours for an extended period of time. Many fractures occur during races, where the public are exposed to upsetting scenes, fuelling debate over the ethics of racing, and research into methods to reduce fracture incidence. Horseracing is a global industry, and it is probably highly unlikely that a curtailment of the number or type of race meetings would be enforced on welfare grounds alone, thus the racing industry and veterinarians have a responsibility to

make racing as safe as possible in its current form. Studies into the causes and risk factors for fracture and other injuries have been appearing in veterinary literature for decades. Unfortunately, sound generic advice on injury avoidance for all racing stakeholders has not been forthcoming, for a number of reasons. Not only do published studies differ in the characteristics of the study population (e.g. age, sex distribution, breed), the injury studied, and the type of activities involved (e.g. racing, training), but significant relationships between risk factors and outcomes have been difficult to interpret and occasionally antagonistic between studies. This library of information is undoubtedly useful in the study of equine disease, and helps us to better understand the risks involved; however, it has not been possible to extrapolate important findings confidently and communicate these to trainers and owners. It also remains to be seen whether such advice would be actively implemented by stakeholders, given the heritage of the sport and the tendency for skills and knowledge on horse training to be passed down the generations within families.

In this study, we identified that fracture, or distal limb fracture, is significantly heritable; a novel finding in these populations. In the Hong Kong population, the heritability of all types of fracture, whether during racing or training, accounts for up to 7% of the total variation. Due to modelling difficulties and the nature of the data available, the 'true' heritability of fracture may be considerably greater than this. It has been shown previously that the nature of fractures differs between those sustained in racing or in training activities (Verheyen et al., 2006; Parkin, 2008). Also, it is unlikely that monotonic fractures have a significant genetic component to their aetiology, aside from the possible role of heritable behavioural traits. If the case definition used in Hong Kong could be made more specific, heritability could possibly be a great deal higher. Nevertheless, the findings reported here are noteworthy, and provide good evidence that further work in this area is warranted.

Fractures of the distal limb, sustained at the racecourse, were found to be significantly heritable in this study. Analyses using BHA data showed that the heritable component of variation may be as large as 23%, which is high for a binary trait. Such fractures are common in racing, and many are hypothesised to be caused by accumulated microdamage and fatigue (Parkin et al., 2006; Kristoffersen et al., 2010a; Whitton et al., 2010). Certain environmental exposures make these fractures more likely (e.g. racing over jumps), but genetic inheritance of risk alleles also plays an important part. These 'fracture genes' may cause a susceptibility to accumulation of microdamage, a tendency for such damage to fail to heal in a timely fashion, or some other complicated biomechanical fault that renders

bone susceptible to catastrophic breakdown. Underlying mechanisms will remain unknown until molecular studies can identify the genes involved. This section of the study was not able to include the time spent at risk of fracture away from the racecourse. As mentioned, many fractures occur during training, and a proportion occur during rest periods. The heritable component of fracture risk may not be restricted to 'racing fractures', or there may be multiple types of heritable fracture with different underlying causative pathways. Future studies with more specific case definitions and accurate diagnoses are warranted to answer these questions.

The results from the BHA and HKJC data are not directly comparable, due to the aforementioned differences in case definitions. However, both serve to increase the index of suspicion that a proportion of our efforts on injury reduction should be focused upon genetic studies in future. Not only do we stand to better understand the nature of fractures in Thoroughbreds, but the horse may serve as a useful model species for human fracture research, thus fostering cross-species research links and bolstering the concept of 'one health'.

Tendon and ligament injuries are very common in racehorses, and necessitate extended rest periods or retirement (Hill et al., 2001; O'Meara et al., 2010). Re-injury is common, and interventional therapies differ widely in efficacy. Certain horse characteristics have been repeatedly linked to increased risk, such as male sex, and many exercise-, course- and race-level risk factors have been identified (Brommer et al., 2004; Kasashima et al., 2004; Lam et al., 2007b). Tendon injuries have received more attention by the research community than ligament injuries, with several high-quality studies being published in recent times (Brommer et al., 2004; O'Meara et al., 2010; Thorpe et al., 2010). The SDFT is by far the most common tendon injury diagnosed in racehorses, therefore most studies focus on this condition specifically (Brommer et al., 2004; Kasashima et al., 2004). A significantly heritable component to SDFT injury was reported by Oki et al. (2008), but to the author's knowledge no further studies of the genetics of Thoroughbred SDFT injury followed. The findings reported here show that both 'all tendon injuries' in Hong Kong, and SDFT injuries in the UK, are significantly heritable, at around 19% and 24%, respectively. In both jurisdictions the EBVs for tendon injury showed a significant trend towards reduction in genetic risk. As has been found previously, exercise history plays a key part in the environmental variation of risk, but interpretation of the relationship is not straightforward. The magnitude of heritability of tendon injury seems to be high for a binary trait, and

certainly of sufficient magnitude that selective breeding could affect a reduction in future genetic risk. Given the difficulties in interpretation of how environmental risk factors affect the risk of tendon injury for individual horses, interventions that target the underlying genetic susceptibility to SDFT injury could provide a vital alternative for stakeholders. The use of published EBVs (or GEBVs based on GS) for careful mate selection is possible, and following future molecular investigations, commercial tests could be used to identify individuals carrying risk alleles. The integration of sound advice on training and management, with knowledge of a horse's individual genetic risk, should be the future of control of tendon injuries in racehorses. High re-injury rates, coupled with often ineffective treatment modalities make it imperative that this information is used to avoid injury in the first place.

A similar stance should be taken with regard to ligament injuries. Although somewhat less prevalent in many jurisdictions, injuries to ligaments are also serious and career-limiting. They do not necessarily have the same risk factors as tendon injuries, and are genetically different from tendon injuries as seen in this study. For this reason, ligament injuries should be considered as important conditions in their own right, and not coupled with tendon injuries in future genetic or environmental studies. The heritability of injuries to the suspensory ligament in Hong Kong Thoroughbreds was highly significant, and the highest estimate was around 16%. Should EBVs for important diseases be considered for use in Thoroughbred breeding schemes, indices that combine information on many conditions are warranted, in a similar way to the indices for cattle breeding that are used currently (Pritchard et al., 2012). Thus breeders would be free to choose the combination of genetic attributes in a stallion that would complement their mares, to influence the conditions most important to their enterprise. Interestingly, a positive genetic correlation between fracture and injury to the suspensory ligament of about 60% was identified here. This finding gives hope that these conditions could be targeted together in breeding schemes, but may also help to increase our knowledge of the genetic pathways underlying the risks of these conditions. Studies of the genes involved are warranted. For both tendon and ligament injuries, future genetic studies using more specific case definitions, and diagnostic imaging, are required to better understand the underlying genetic implications. Whether these studies are forthcoming or not, the findings presented here give further evidence that veterinarians and trainers should be vigilant, particularly with older male horses competing in jump races, as these appear to be most at risk of tendon or ligament injury.

Pathology of articular joints is common in Thoroughbreds, and can stem from many causes (McIlwraith and Vachon, 1988; Jeffcott, 1997; Santschi et al., 1998; Brommer et al., 2004; Neundorf et al., 2010; Corbin et al., 2012; Reed et al., 2012). Trauma, fatigue, infection and developmental disorders, amongst others, can all lead to alterations in the normal anatomy and physiology of the cartilaginous and/or osseous joint tissues. Identification of the inciting cause, or most likely aetiology, is often possible via the use of sophisticated imaging technologies such as magnetic resonance imaging and computed tomography. In some cases, the pathology identified cannot be assigned to a specific underlying cause. Without the use of such diagnostic imaging technologies, the clinician is restricted to evaluation of clinical signs, perineural or intra-articular anaesthesia and careful consideration of the history to make tentative diagnoses. Swelling, heat, pain, reduced range of motion and lameness may all ensue, but identification of the joints involved is not always easy. In the HKJC data used here, many horses were diagnosed as having osteoarthritis or degenerative joint disease. Due to the range of differing pathological mechanisms involved both within and between these groups, and the absence of information on the diagnostic aids used, the case definition of joint disease used in this study may have suffered from low specificity and sensitivity. Had heritability of joint disease been found to be moderate or large, and unequivocal in significance, the use of selective breeding based on EBVs could have been considered. Unfortunately, this was not the case, probably because of the lack of certainty of the causes of joint disease phenotypes. Due to the prevalence of joint problems in racehorses, it would be beneficial to repeat heritability analyses with larger case numbers and specific case definitions, in order to identify the nature of the joint pathologies that gave rise to the findings reported here, and to provide targets for future breeding strategies.

Exercise-induced pulmonary haemorrhage occurs in racehorses worldwide. A number of theoretical aetiologies have been posited to explain the condition, but none has been shown to be dominant. It has been generally accepted that the prevalence of EIPH is high in all types of racing, in all breeds of racehorse and across all racing jurisdictions (Lapointe et al., 1994; Takahashi et al., 2001; Newton et al., 2005). The association of EIPH and related epistaxis have been accepted by most as being related to poor performance, thus interventions such as the pre-race use of furosemide, and research into causative mechanisms and risk factors have been implemented by many groups. Furosemide is prohibited for use during racing in the UK and Hong Kong, due to its performance-

enhancing effects, as well as to maintain the good public image of the sport. Whether EIPH and epistaxis compromise horse welfare is debatable, but studies have shown a link between severe pulmonary haemorrhage and sudden death, therefore reduction of EIPH incidence is of veterinary interest (Johnson et al., 1994; McKee, 1995). It is unknown at the time of writing whether horses that die as a result of such severe haemorrhage constitute a different genetic group to the vast majority that experience a lesser degree of EIPH. Only one study in recent times has estimated the heritability of epistaxis as related to EIPH, and this study found a highly significant heritability of 0.23 to 0.4 (Weideman et al., 2004). The results presented here agree with that study. EIPH/epistaxis diagnosed during racing or training in Hong Kong, partly by visual inspection without tracheobronchoscopy in all cases, was found to have a heritability of up to 28%. Epistaxis at the racecourse in the UK was 26% heritable, whether horses were jump racers, flat racers, or mixed racing horses. Diagnosis of epistaxis relies on trainers or veterinarians noticing what may be a small volume of blood at the nostrils following exercise, which may be difficult in the busy atmosphere of the racecourse, or if bleeding occurs after some delay. These figures give strong evidence that EIPH/epistaxis is a highly heritable trait, which may be manipulated using selective breeding, should the racing industry adopt an attitude conducive to change. Compelling studies of the possible benefits to welfare, performance and racing economics may be required before substantial changes in breeding strategy are adopted by the industry wholesale. Furosemide ('Lasix') is widely used in the USA to prevent epistaxis during racing. This drug has been shown to have performance-enhancing effects (Gross 1999). Horses that were deemed likely to 'bleed' during racing may have been selected for breeding over those that were not likely to 'bleed', as their performance may have been improved by the administration of furosemide. This selective breeding could have led to an increase in the prevalence of EIPH in the population over time. Despite this, no trend in EBVs for EIPH/epistaxis was found over the study period. This may be due to horses being imported from all over the world to Hong Kong, including from the UK where the use of furosemide is prohibited.

Historically, animal breeding has been implemented to improve production in farm animal species. More recently, inclusion of factors associated with animal welfare and longevity have been incorporated, to comply with animal welfare legislation and to improve animal health and wellbeing in a more holistic fashion (Pritchard et al., 2012). It has been noted that inclusion of longevity traits in animal breeding improves profitability for farmers in the longer term, despite possible reductions in production in the immediate short term

(Pritchard et al., 2012). Bull breeding indices combine a number of production and welfare traits into a profile for each animal, which allows farmers to choose animals that best complement their existing stock. The Profitable Lifetime Index (£PLI) is used in the UK to identify bulls whose offspring are likely to generate more income for the farmer than lower ranked bulls. The £PLI incorporates many heritable traits related to production (e.g. kilograms of milk), along with traits for health and longevity, to enable farmers to maximise income whilst safeguarding the welfare of cattle. Negative genetic correlations between production and health traits have been found in dairy cattle in previous studies, such as between milk yield and resistance to mastitis. Both are included in bull indices, but weighted according to their socioeconomic importance for the industry so that a balance can be struck between lost profits and improvements in animal welfare (Oltenacu et al., 2010). In a similar fashion, it would be of greater benefit to the Thoroughbred breeding industry on the whole if breeding indices combined information on many disease, performance and longevity traits, rather than focusing on individual conditions. This approach may help to avoid the pitfalls associated with strong selection of correlated traits, whereby unintended negative consequences may ensue. It is also important that the goals of horse breeding be made as desirable to stakeholders as possible, by incorporation of factors that will improve returns and/or directly improve performance. In an ideal world, improvement in horse performance and longevity would be closely coupled with good health and welfare, such that veterinary and economic interests would be perfectly aligned. In this study, we investigated three variables associated with performance or durability, to identify whether they were heritable themselves, and to study their associations with important disease conditions. These performance/durability traits would be deemed useful if they were heritable, and negatively genetically correlated with binary diseases. Career length, number of career starts, and winnings were all moderately or highly heritable in both study populations. Some showed favourable genetic correlations with diseases, and some showed deleterious correlations. Overall, none were perfectly aligned with horse health, and all carry caveats to their interpretation. More sophisticated measures of performance or longevity may be investigated in future studies, and their relationships to horse health in a wider context investigated, to elucidate more clearly whether this approach to breeding racehorses could be useful.

The data used in the analyses of horse health, longevity and durability in this study were a valuable resource. The quality of these data enabled preliminary analyses of heritability, and allowed the overall aims of this study to be met. Future analyses of the diseases

studied here would benefit from the use of more specific diagnoses with, in some cases, larger sample sizes and more complete pedigrees to maximise heritability estimate accuracy and avoid computational difficulties. Such investigations, alongside molecular genetic studies, could yield further information about the fundamental nature of each disease entity. The findings reported here suggest that further work to define and analyse the heritability of more holistic traits could prove more fruitful for the practical manipulation of the Thoroughbred horse population, as opposed to focusing on individual diseases alone. In short, the heritability analyses presented here provide the first steps towards better understanding of the nature of many important diseases in Thoroughbred racehorses, and serve as a starting point from which future work could lead to the production of practical advice for stakeholders on how to make genetic improvements to racehorse populations. Should this work influence the discovery of genetic markers for any of the diseases featured, the information contained in this thesis should help the end user understand what a positive test result for such a marker may mean in terms of the risk of the disease outcome. In reality, discovery of an individual marker for disease, or even a panel of such markers, may not mean a great deal in terms of the likelihood of developing the outcome in an individual. However, this does not underplay the importance of the development of such markers, as it will always be of interest to know if an individual is at even slightly increased risk so that their environment can be duly modified in the hope of avoiding the deleterious outcome. The size of the effect of each marker must always be taken into account and emphasised when choosing whether or not to subject animals to testing.

Appendix

A. Dictionaries used in contents analysis of HKJC data. Entries in bold type are category names, and entries in normal font are words or phrases pertaining to that category. * indicates a word fragment.

a) Fracture dictionary.

Fracture	**Fracture – Carpus**	**Fracture – MC/T3 Diaphysis**
BROKE_DOWN	C2*	CANNON
BREAKDOWN	CARP*	DIAPHY*
BROKEN_DOWN	CAPAL	CN
CATASTROPHIC	CARRAL	
CHIP		**Fracture – Tibia**
CN	**Fracture – Sesamoid**	TIB*
COMMINUATED	SEGAMOID*	
COMMINUTED	SESAMOID*	**Fracture – Patella**
COMPOUND		PATELLA
COPMMINUTED	**Fracture – Condyle of MC/T3**	
FACTURE	CONDYL*	**Fracture – Navicular**
FRACTURE*		NAVICULAR
FRACUTRED	**Fracture – Skull**	
FRACUTRE	SKULL	**Fracture – Splint**
SLAB*		SPLINT
	Fracture – Pelvic	
Fracture – Other Sites	COXAE	**Fracture – Scapula**
FETL*	ILI*	SCAPULA
FTK	ILLAL_WING	
HOCK	PELV*	**Fracture – Ulna**
KNEE*	ILLIAL_WING	OLECRANON
STIFL*		
SHOULDER	**Fracture – Humerus**	**Fracture – Distal Phalanx**
	HUMER*	PEDAL
Fracture – Proximal Phalanx		
P1*	**Fracture – Radius**	
PASTERN	RADI*	

b) Tendon condition dictionary.

Tendon	**Ligament**	**Ligament - Sesamoidean**
DIGITAL_FLEXOR	DEMITIS	SESAMOIDEAN
FLEXOR	DESMITIS	
PERITENDINUOUS	LEGAMENT	**Ligament - Check**
SDF*	LIGAMENT*	CHECK
SUPERFICIAL_DIGITAL	SL	
TEDON*	SUPSENS*	**Ligament - Illiac**
TEND*	SUSPEN*	ILLIAC
TENO*	SUSPL	
TN_INJURY		**Ligament - Suspensory**
WINDGALL	**Ligament - Collateral**	SUPENS*
	COLLATERAL	SUSPEN*
		SUPSEN*

c) EIPH and bleeding dictionary

Bleeding	**Bleed- Once**	**Bleed - Twice**
BLED	1ST	SECOND
BLEED*	FIRST	2ND
BLLEDING	ONCE	
BLOOD	ONEC	**Bleed - Three Times**
EIPH*		3RD
EPIH*		CHRONIC

d) Poor performance dictionary

Poor Performance	LOST_ABILITY	PERFORMANNCE
ABILITY	LOST_INTEREST	PERFORMS
BARRIER	LOW_RATING	POOR_FORELIMB_ACTION
CL	NON_COMPETITIVE*	POOR_FORM
CL6	NON_STARTER*	POOR_HIND_LIMB_ACTION
CL6*	NON_STRATERS	POOR_PERFORMANCE
CLASS	NOT_GOOD_ENOUGH	NO_ABILITY
CLASS_4	NOT_INTERESTED	PROFORMANCE
CLASS_5	NO_GOOD	RATED
CLASS_6	NO_LONGER	RATE_BELOW
FORM	NO_ROOM_FOR_IMPROVEMENT	SPEEDY_CUTTING
GATE*	PEAK_LEVEL	STIFFNESS*
INABILITY	PERFMORANCE	UNCOORDINATED
INTEREST	PERFORMAENCE	USELESS
LACK_OF _ABILITY	PERFORMAMCE	UNSUITABLE
LIMITED	PERFORMANACE	
LOSS_OF_FORM	PERFORMANCE*	

e) Musculoskeletal condition dictionary

Musculoskeletal	FETLOCK*	LEG_PROBLEM
ACCIDENT	FOOT	LUXATED
BAD_LEG	GULTEAL	LUXATION
ATROPHY	HEAD	MUSCLE
BRUISING	INJURY	PASTERN*
BURSA	JOINT	SACROILIA*
CAPITIS	KNEE	SESAMOIDITIS
CARPITIS	LAME	SORE_SHIN
CHONDRITIS	LAMENESS	SPLINT
COFFIN	LAMENSS	SUBLUXATION
CRACK	LAMINITIS	TRAUMA
DISLOCATED	LAMNITIS	UNSOUND
DISLOCATION	LEGS	VERTEBRAE

f) DJD and OA dictionary

DJD	DEGEMERATIVE	OSTEOARTHRITIS
ARTHRITI*	DEGENERATIVE_JOINT	OSTEOURTHRITIS
ASTEOARTHRITIS		OSTEOPHYTE
CHRONIC_DEGENERATIVE_DISEASE	**OA**	
DJD	OA	

g) Medical condition dictionary

Medical	EUTHANASED	NEUROLOGICAL
ABDOMINAL	EVISCERATION	NOSTRILS
ANAEMIA	EYE	NOSTRIALS
ANAPHYLACTIC	FIBRILLATION	PAIN
ANAPHYLAXIC	FIBROSIS	PALATE*
AORTIC	FRACTIOUS	PARALYSIS
ARRHYTHMIA	G/A	PEDAL
ARYTENOID*	GA	PEMPHIGUS
ATAXI*	GASTRIC	PENIS
ATRIAL	GLUTEAL	PERITONITIS
AUTOIMMUNE	HAEMATURIA	PHARYNGEAL
BEHAVIONAL	HAEMOLYTIC	PLEUROPNEUMONIA
BEHAVIOR	HAIR	POOR_CONDITION
BEHAVIOUR	HEALTH	POOR_HEALTH
BLIND	HEART	PYREXIA
BLINDNESS	HEAT_DISTRESS	RECALCITRANT
BOX_WALKER	HORMONE	RECTA*
BRAIN	HUMANITARIAN	RESPIRATORY
BREATHING	HYGROMA	RHYTHM
BRONCHOPNEUMONIC	IMPACTION	SCOUR*
CAECAL	INAPPETENCE	SEPSIS
CAECUM	INFECTED	SHOCK
CARDIO	INFECTION	SKIN
CELLUCITIS	INTESTINAL	STALLS
CELLULITIS	INTESTINE	STOMACH
COLIC	INTRACTABLE	SUDDEN_DEATH
COLITIS	IRREGULARITY	TEMPER*
COLLAPSE*	JUGULAR	THORAX
CONGESTIVE	LACERATIONS	THROAT
CORNEAL	LACEATEONS	TOXIC
DEAD	LARYNGAL	TRACHEAL
DEATH	LARYNGEAL	TUMOUR
DERMATITIS	LARYNGOAL	TWISTED
DIAPHRAGMATIC	LARYNGOPLASTY	TYPHLITIS
DIARRHOEA	LAYNGEAL	ULCER*
DIED	LESIONS	UNRULY
DIFFICULT	LEUCOCYTOSIS	WEIGHT_LOSS
DYSPHAGIA	LUNG	VIRUS
ENDOTOXAEMIA	MENTAL	WOBBLER
ENDOTOXIC	MURMUR	WHITE_CELL
EPIGLOTTIC	NERVOUS	NEUROLOGICAL

h) Management decision dictionary

Management	HOLD	TRAINER
ADVICE_TO_RETIRE	NO_BID*	TRAINER'S_ADVICE*
FINANCIAL	RESIDENCY	TRANSFER
LOT	STEWARDS_REQUEST	STUD
MEMBERSHIP	SYNDICATE	UNRACED

B. Publications based on this work

The Veterinary Journal, June 5th, 2013, pii: S1090-0233(13)00200-1. doi: 10.1016/j.tvjl.2013.05.002

PRELIMINARY GENETIC ANALYSES OF IMPORTANT MUSCULOSKELETAL CONDITIONS OF THOROUGHBRED RACEHORSES IN HONG KONG

Welsh CE, Lewis TW, Blott SC, Mellor DJ, Lam KH, Stewart BD, Parkin TDH

The Veterinary Journal, Submitted October 1st 2013, with reviewers at time of printing

ESTIMATES OF GENETIC PARAMETERS OF DISTAL LIMB FRACTURE AND SUPERFICIAL DIGITAL FLEXOR TENDON INJURY IN UK THOROUGHBRED RACEHORSES

Welsh CE, Lewis TW, Blott SC, Mellor DJ, Stirk A, Parkin TDH

References

Akesson, M., Bensch, S., Hasselquist, D., Tarka, M., Hansson, B., 2008. Estimating heritabilities and genetic correlations: comparing the 'animal model' with parent-offspring regression using data from a natural population. PLoS One 3, e1739.

Anthenill, L.A., Gardner, I.A., Pool, R.R., Garcia, T.C., Stover, S.M., 2010. Comparison of macrostructural and microstructural bone features in Thoroughbred racehorses with and without midbody fracture of the proximal sesamoid bone. American Journal of Veterinary Research 71, 755-765.

Bailey, C.J., Reid, S.W.J., Hodgson, D.R., Bourke, J.M., Rose, R.J., 1998. Flat, hurdle and steeple racing: risk factors for musculoskeletal injury. Equine Veterinary Journal 30, 498-503.

Bakhtiari, J., Kashan, N.E.J., 2009. Estimation of genetic parameters of racing performance in Iranian Thoroughbred horses. Livestock Science 120, 151-157.

Bannasch, D., 2008. Genetic Testing and the Future of Equine Genomics. Journal of Equine Veterinary Science 28, 645-649.

Beisser, A.L., McClure, S., Wang, C., Soring, K., Garrison, R., Peckham, B., 2011. Evaluation of catastrophic musculoskeletal injuries in Thoroughbreds and Quarter Horses at three Midwestern racetracks. Journal of the American Veterinary Medical Association 239, 1236-1241.

BHA, 2009. Economic impact of British racing.

Binns, M.M., Boehler, D.A., Bailey, E., Lear, T.L., Cardwell, J.M., Lambert, D.H., 2012. Inbreeding in the Thoroughbred horse. Anim Genet 43, 340-342.

Bishop, S.C., Woolliams, J.A., 2010. On the Genetic Interpretation of Disease Data. PLoS One 5.

Boden, L.A., Anderson, G.A., Charles, J.A., Morgan, K.L., Morton, J.M., Parkin, T.D.H., Clarke, A.F., Slocombe, R.F., 2007. Risk factors for Thoroughbred racehorse fatality in flat starts in Victoria, Australia (1989-2004). Equine Veterinary Journal 39, 430-437.

Boden, L.A., Charles, J.A., Slocombe, R.F., Sandy, J.R., Finnin, P.J., Morton, J.M., Clarke, A.F., 2005. Sudden death in racing Thoroughbreds in Victoria, Australia. Equine Vet J 37, 269-271.

Bolwell, C.F., Rogers, C.W., French, N.P., Firth, E.C., 2012. Risk factors for interruptions to training occurring before the first trial start of 2-year-old Thoroughbred racehorses. New Zealand Veterinary Journal 60, 241-246.

Boyde, A., Firth, E.C., 2005. Musculoskeletal responses of 2-year-old Thoroughbred horses to early training. 8. Quantitative back-scattered electron scanning electron microscopy and

confocal fluorescence microscopy of the epiphysis of the third metacarpal bone. New Zealand Veterinary Journal 53, 123-132.

Brommer, H., Brama, P.A., Barneveld, A., van Weeren, P.R., 2004. Differences in the topographical distribution of articular cartilage degeneration between equine metacarpo- and metatarsophalangeal joints. Equine Vet J 36, 506-510.

Brosnahan, M.M., Brooks, S.A., Antczak, D.F., 2010. Equine clinical genomics: A clinician's primer. Equine Veterinary Journal 42, 658-670.

Bugislaus, A.E., Roehe, R., Willms, F., Kalm, E., 2006. The use of a random regression model to account for change in racing speed of German trotters with increasing age. Journal of Animal Breeding and Genetics 123, 239-246.

Buxadera, A.M., da Mota, M.D.S., 2008. Variance component estimations for race performance of thoroughbred horses in Brazil by random regression model. Livestock Science 117, 298-307.

Cogger, N., Evans, D.L., Hodgson, D.R., Reid, S.W., Perkins, N., 2008. Incidence rate of musculoskeletal injuries and determinants of time to recovery in young Australian Thoroughbred racehorses. Australian Veterinary Journal 86, 473-480.

Cogger, N., Perkins, N., Hodgson, D.R., Reid, S.W.J., Evans, D.L., 2006. Risk factors for musculoskeletal injuries in 2-year-old Thoroughbred racehorses. Preventive Veterinary Medicine 74, 36-43.

Cohen, N.D., Berry, S.M., Peloso, J.G., Mundy, G.D., Howard, I.C., 2000. Association of high-speed exercise with racing injury in Thoroughbreds. Journal of the American Veterinary Medical Association 216, 1273-1278.

Corbin, L.J., Blott, S.C., Swinburne, J.E., Sibbons, C., Fox-Clipsham, L.Y., Helwegen, M., Parkin, T.D., Newton, J.R., Bramlage, L.R., McIlwraith, C.W., Bishop, S.C., Woolliams, J.A., Vaudin, M., 2012. A genome-wide association study of osteochondritis dissecans in the Thoroughbred. Mamm Genome 23, 294-303.

Corbin, L.J., Blott, S.C., Swinburne, J.E., Vaudin, M., Bishop, S.C., Woolliams, J.A., 2010. Linkage disequilibrium and historical effective population size in the Thoroughbred horse. Anim Genet 41 Suppl 2, 8-15.

Costa, M.F., Ronchi, F.A., Ivanow, A., Carmona, A.K., Casarini, D., Slocombe, R.F., 2012. Association between circulating angiotensin-converting enzyme and exercise-induced pulmonary haemorrhage in Thoroughbred racehorses. Res Vet Sci 93, 993-994.

Cunningham, E.P., Dooley, J.J., Splan, R.K., Bradley, D.G., 2001. Microsatellite diversity, pedigree relatedness and the contributions of founder lineages to thoroughbred horses. Animal Genetics 32, 360-364.

Daetwyler, H.D., Villanueva, B., Bijma, P., Woolliams, J.A., 2007. Inbreeding in genome-wide selection. J Anim Breed Genet 124, 369-376.

Davidson, E., 2003. Clinical recognition of stress-related bone injury in racehorses. Clinical Techniques in Equine Practice 2, 296-311.

Davidson, E.J., Martin, B.B., Boston, R.C., Parente, E.J., 2011. Exercising upper respiratory videoendoscopic evaluation of 100 nonracing performance horses with abnormal respiratory noise and/or poor performance. Equine Vet J 43, 3-8.

Dempster, E.R., Lerner, I.M., 1950. Heritability of Threshold Characters. Genetics 35, 212-236.

Dohoo, I.R., Martin, W., Stryhn, H., 2003. Veterinary Epidemiologic Research. University of Prince Edward Island, ISBN: 0919013414

Dudhia, J., Scott, C.M., Draper, E.R., Heinegard, D., Pitsillides, A.A., Smith, R.K., 2007. Aging enhances a mechanically-induced reduction in tendon strength by an active process involving matrix metalloproteinase activity. Aging Cell 6, 547-556.

Duncan, E.L., Danoy, P., Kemp, J.P., Leo, P.J., McCloskey, E., Nicholson, G.C., Eastell, R., Prince, R.L., Eisman, J.A., Jones, G., Sambrook, P.N., Reid, I.R., Dennison, E.M., Wark, J., Richards, J.B., Uitterlinden, A.G., Spector, T.D., Esapa, C., Cox, R.D., Brown, S.D., Thakker, R.V., Addison, K.A., Bradbury, L.A., Center, J.R., Cooper, C., Cremin, C., Estrada, K., Felsenberg, D., Gluer, C.C., Hadler, J., Henry, M.J., Hofman, A., Kotowicz, M.A., Makovey, J., Nguyen, S.C., Nguyen, T.V., Pasco, J.A., Pryce, K., Reid, D.M., Rivadeneira, F., Roux, C., Stefansson, K., Styrkarsdottir, U., Thorleifsson, G., Tichawangana, R., Evans, D.M., Brown, M.A., 2011. Genome-wide association study using extreme truncate selection identifies novel genes affecting bone mineral density and fracture risk. PLoS Genet 7, e1001372.

Dyson, P.K., Jackson, B.F., Pfeiffer, D.U., Price, J.S., 2008. Days lost from training by two- and three-year-old Thoroughbred horses: a survey of seven UK training yards. Equine Vet J 40, 650-657.

Ekiz, B., Kocak, O., 2007. Estimates of genetic parameters for racing times of thoroughbred horses. Turkish Journal of Veterinary & Animal Sciences 31, 1-5.

Ely, E.R., Avella, C.S., Price, J.S., Smith, R.K.W., Wood, J.L.N., Verheyen, K.L.P., 2009. Descriptive epidemiology of fracture, tendon and suspensory ligament injuries in National Hunt racehorses in training. Equine Veterinary Journal 41, 372-378.

Ely, E.R., Verheyen, K.L.P., Wood, J.L.N., 2004. Fractures and tendon injuries in National Hunt horses in training in the UK: a pilot study. Equine Veterinary Journal 36, 365-367.

Entwistle, R., Sammons, S., Hazelwood, S., Fyhrie, D., Stover, S., 2008. Exercise history and remodelling stress fracture are related to cortical bone ultimate strength (P264). Engineering of Sport 7, Vol 2, 613-621.

Estberg, L., Stover, S.M., Gardner, I.A., Johnson, B.J., Case, J.T., Ardans, A., Read, D.H., Anderson, M.L., Barr, B.C., Daft, B.M., Kinde, H., Moore, J., Stoltz, J., Woods, L.W., 1996. Fatal musculoskeletal injuries incurred during racing and training in thoroughbreds. Journal of the American Veterinary Medical Association 208, 92-96.

Estberg, L., Stover, S.M., Gardner, I.A., Johnson, B.J., Jack, R.A., Case, J.T., Ardans, A., Read, D.H., Anderson, M.L., Barr, B.C., Daft, B.M., Kinde, H., Moore, J., Stoltz, J., Woods, L., 1998. Relationship between race start characteristics and risk of catastrophic injury in thoroughbreds: 78 cases (1992). Journal of the American Veterinary Medical Association 212, 544-+.

Falconer, D.S., Mackay, T.F.C., 1996. Introduction to Quantitative Genetics 4th Edition. Oliver & Boyd.

Gianola, D., 1980. A method of sire evaluation for dichotomies. J Anim Sci 51, 1266-1271.

Gianola, D., 1982. Theory and Analysis of Threshold Characters. J Anim Sci 54, 1079-1096.

Gu, J.J., Orr, N., Park, S.D., Katz, L.M., Sulimova, G., MacHugh, D.E., Hill, E.W., 2009. A Genome Scan for Positive Selection in Thoroughbred Horses. PLoS One 4.

Haberland, A.M., Konig von Borstel, U., Simianer, H., Konig, S., 2012. Integration of genomic information into sport horse breeding programs for optimization of accuracy of selection. Animal 6, 1369-1376.

Hamann, H., Distl, O., 2008. Genetic variability in Hanoverian warmblood horses using pedigree analysis. J Anim Sci 86, 1503-1513.

Hayes, B.J., Bowman, P.J., Chamberlain, A.J., Goddard, M.E., 2009. Invited review: Genomic selection in dairy cattle: progress and challenges. Journal of Dairy Science 92, 433-443.

Hernandez, J., Hawkins, D.L., Scollay, M.C., 2001. Race-start characteristics and risk of catastrophic musculoskeletal injury in Thoroughbred racehorses. Journal of the American Veterinary Medical Association 218, 83-86.

Hill, A.E., Stover, S.M., Gardner, I.A., Kane, A.J., Whitcomb, M.B., Emerson, A.G., 2001. Risk factors for and outcomes of noncatastrophic suspensory apparatus injury in Thoroughbred racehorses. Journal of the American Veterinary Medical Association 218, 1136-1144.

Hill, E.W., Fonseca, R.G., McGivney, B.A., Gu, J.J., MacHugh, D.E., Katz, L.M., 2012. MSTN genotype (g.66493737C/T) association with speed indices in Thoroughbred racehorses. Journal of Applied Physiology 112, 86-90.

Hill, E.W., McGivney, B.A., Gu, J., Whiston, R., Machugh, D.E., 2010. A genome-wide SNP-association study confirms a sequence variant (g.66493737C>T) in the equine myostatin (MSTN) gene as the most powerful predictor of optimum racing distance for Thoroughbred racehorses. BMC Genomics 11, 552.

Hinchcliff, K.W., Jackson, M.A., Morley, P.S., Brown, J.A., Dredge, A.E., O'Callaghan, P.A., McCaffrey, J.P., Slocombe, R.E., Clarke, A.E., 2005. Association between exercise-induced pulmonary hemorrhage and performance in Thoroughbred racehorses. Journal of the American Veterinary Medical Association 227, 768-774.

Hinchcliff, K.W., Morley, P.S., Guthrie, A.J., 2009. Efficacy of furosemide for prevention of exercise-induced pulmonary hemorrhage in Thoroughbred racehorses. Journal of the American Veterinary Medical Association 235, 76-82.

Hinchcliff, K.W., Morley, P.S., Jackson, M.A., Brown, J.A., Dredge, A.F., O'Callaghan, P.A., McCaffrey, J.P., Slocombe, R.F., Clarke, A.F., 2010. Risk factors for exercise-induced pulmonary haemorrhage in Thoroughbred racehorses. Equine Veterinary Journal 42, 228-234.

HKJC 2010. Annual Report 2009/2010 (The Hong Kong Jockey Club).

HKJC 2011. 2010/11 Season End Results, http://corporate.hkjc.com/corporate/operation/english/10-11-results.aspx.

Hoque, M.A., Kadowaki, H., Shibata, T., Oikawa, T., Suzuki, K., 2007. Genetic parameters for measures of the efficiency of gain of boars and the genetic relationships with its component traits in Duroc pigs. Journal of Animal Science 85, 1873-1879.

Jeffcott, L., 1997. Osteochondrosis in horses. In Practice 19, 64-71.

Johnson, B.J., Stover, S.M., Daft, B.M., Kinde, H., Read, D.H., Barr, B.C., Anderson, M., Moore, J., Woods, L., Stoltz, J., Blanchard, P., 1994. Causes of Death in Racehorses over a 2-Year Period. Equine Veterinary Journal 26, 327-330.

Jonsson, L., Dalin, G., Egenvall, A., Nasholm, A., Roepstorff, L., Philipsson, J., 2011. Equine hospital data as a source for study of prevalence and heritability of osteochondrosis and palmar/plantar osseous fragments of Swedish Warmblood horses. Equine Vet J 43, 695-700.

Kalisiak, O., 2012. Parameters influencing prevalence and outcome of tendonitis in Thoroughbred and Arabian racehorses. Polish Journal of Veterinary Sciences 15, 111-118.

Kasashima, Y., Takahashi, T., Smith, R.K.W., Goodship, A.E., Kuwano, A., Ueno, T., Hirano, S., 2004. Prevalence of superficial digital flexor tendonitis and suspensory desmitis in Japanese Thoroughbred flat racehorses in 1999. Equine Veterinary Journal 36, 346-350.

Kristoffersen, M., Hetzel, U., Parkin, T.D., Singer, E.R., 2010a. Are bi-axial proximal sesamoid bone fractures in the British Thoroughbred racehorse a bone fatigue related fracture? A histological study. Vet Comp Orthop Traumatol 23, 336-342.

Kristoffersen, M., Parkin, T.D., Singer, E.R., 2010b. Catastrophic biaxial proximal sesamoid bone fractures in UK Thoroughbred races (1999-2004): horse characteristics and racing history. Equine Vet J 42, 420-424.

Lam, K.H., Parkin, T.D.H., Riggs, C.M., Morgan, K.L., 2007a. Descriptive analysis of retirement of Thoroughbred racehorses due to tendon injuries at the Hong Kong Jockey Club (1992-2004). Equine Veterinary Journal 39, 143-148.

Lam, K.K.H., Parkin, T.D.H., Riggs, C.M., Morgan, K.L., 2007b. Evaluation of detailed training data to identify risk factors for retirement because of tendon injuries in Thoroughbred racehorses. American Journal of Veterinary Research 68, 1188-1197.

Lanyon, L.E., 1987. Functional strain in bone tissue as an objective, and controlling stimulus for adaptive bone remodelling. J Biomech 20, 1083-1093.

Lapointe, J.M., Vrins, A., McCarvill, E., 1994. A survey of exercise-induced pulmonary haemorrhage in Quebec standardbred racehorses. Equine Vet J 26, 482-485.

Le Jeune, S.S., Macdonald, M.H., Stover, S.M., Taylor, K.T., Gerdes, M., 2003. Biomechanical Investigation of the Association Between Suspensory Ligament Injury and Lateral Condylar Fracture in Thoroughbred Racehorses. Veterinary Surgery 32, 585-597.

Lewis, T.W., Blott, S.C., Woolliams, J.A., 2010. Genetic evaluation of hip score in UK Labrador Retrievers. PLoS One 5, e12797.

Loitz, B.J., Zernicke, R.F., 1992. Strenuous Exercise-Induced Remodeling of Mature Bone - Relationships between Invivo Strains and Bone Mechanics. Journal of Experimental Biology 170, 1-18.

Lyle, C.H., Blissitt, K.J., Kennedy, R.N., McGorum, B.C., Newton, J.R., Parkin, T.D.H., Stirk, A., Boden, L.A., 2012. Risk factors for race-associated sudden death in Thoroughbred racehorses in the UK (2000-2007). Equine Veterinary Journal 44, 459-465.

Mahon, G.A.T., Cunningham, E.P., 1982. Inbreeding and the Inheritance of Fertility in the Thoroughbred Mare. Livestock Production Science 9, 743-754.

Malm, S., Fikse, W.F., Danell, B., Strandberg, E., 2008. Genetic variation and genetic trends in hip and elbow dysplasia in Swedish Rottweiler and Bernese Mountain Dog. J Anim Breed Genet 125, 403-412.

Matos, C.A., Thomas, D.L., Gianola, D., Tempelman, R.J., Young, L.D., 1997. Genetic analysis of discrete reproductive traits in sheep using linear and nonlinear models: I. Estimation of genetic parameters. J Anim Sci 75, 76-87.

McGivney, B.A., Eivers, S.S., MacHugh, D.E., MacLeod, J.N., O'Gorman, G.M., Park, S.D.E., Katz, L.M., Hill, E.W., 2009. Transcriptional adaptations following exercise in Thoroughbred horse skeletal muscle highlights molecular mechanisms that lead to muscle hypertrophy. BMC Genomics 10, 638.

McIlwraith, C.W., Vachon, A., 1988. Review of pathogenesis and treatment of degenerative joint disease. Equine Vet J Suppl, 3-11.

McKee, S.L., 1995. An update on racing fatalities in the UK. Equine Veterinary Education 7, 202-204.

Meuwissen, T.H., Hayes, B.J., Goddard, M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157, 1819-1829.

Meyer, T.S., Fedde, M.R., Gaughan, E.M., Langsetmo, I., Erickson, H.H., 1998. Quantification of exercise-induced pulmonary haemorrhage with bronchoalveolar lavage. Equine Vet J 30, 284-288.

Michelotto, P.V., Jr., Muehlmann, L.A., Zanatta, A.L., Bieberbach, E.W., Kryczyk, M., Fernandes, L.C., Nishiyama, A., 2011. Pulmonary inflammation due to exercise-induced pulmonary haemorrhage in Thoroughbred colts during race training. Vet J 190, e3-6.

Mittmann, E.H., Momke, S., Distl, O., 2010. Whole-genome scan identifies quantitative trait loci for chronic pastern dermatitis in German draft horses. Mamm Genome 21, 95-103.

Mota, M.D.S., Abrahao, A.R., Oliveira, H.N., 2005. Genetic and environmental parameters for racing time at different distances in Brazilian Thoroughbreds. Journal of Animal Breeding and Genetics 122, 393-399.

Neundorf, R.H., Lowerison, M.B., Cruz, A.M., Thomason, J.J., McEwen, B.J., Hurtig, M.B., 2010. Determination of the prevalence and severity of metacarpophalangeal joint osteoarthritis in Thoroughbred racehorses via quantitative macroscopic evaluation. American Journal of Veterinary Research 71, 1284-1293.

Newton, J., 2008. Epidemiology of EIPH. Havermeyer Foundation Monograph Series, 3-7.

Newton, J.R., Rogers, K., Marlin, D.J., Wood, J.L., Williams, R.B., 2005. Risk factors for epistaxis on British racecourses: evidence for locomotory impact-induced trauma contributing to the aetiology of exercise-induced pulmonary haemorrhage. Equine Vet J 37, 402-411.

Nunamaker, D.M., Butterweck, D.M., Provost, M.T., 1990. Fatigue fractures in thoroughbred racehorses: relationships with age, peak bone strain, and training. Journal of Orthopaedic Research 8, 604-611.

Oltenacu, P.A., Broom, D.M., 2010. The impact of genetic selection for increased milk yield on the welfare of dairy cows. Animal Welfare, 19 (S): 39-49.

O'Meara, B., Bladon, B., Parkin, T.D., Fraser, B., Lischer, C.J., 2010. An investigation of the relationship between race performance and superficial digital flexor tendonitis in the Thoroughbred racehorse. Equine Vet J 42, 322-326.

Oikawa, M., Kusunose, R., 2005. Fractures sustained by racehorses in Japan during flat racing with special reference to track condition and racing time. Veterinary Journal 170, 369-374.

Oki, H., Kusunose, R., Nakaoka, H., Nishiura, A., Miyake, T., Sasaki, Y., 2007. Estimation of heritability and genetic correlation for behavioural responses by Gibbs sampling in the Thoroughbred racehorse. J Anim Breed Genet 124, 185-191.

Oki, H., Miyake, T., Kasashima, Y., Sasaki, Y., 2008. Estimation of heritability for superficial digital flexor tendon injury by Gibbs sampling in the Thoroughbred racehorse. Journal of Animal Breeding and Genetics 125, 413-416.

Park, K.D., 2011. Genetic parameters of finish time in Korean Thoroughbred racehorses. Livestock Science 140, 49-54.

Parkin, T., 2008. Epidemiology of Racetrack Injuries in Racehorses. Veterinary Clinics of North America: Equine Practice 24, 1-19.

Parkin, T., Clegg, P., French, N., Proudman, C., Riggs, C., Singer, E., Webbon, P., Morgan, K., 2006. Catastrophic fracture of the lateral condyle of the third metacarpus/metatarsus in UK racehorses – fracture descriptions and pre-existing pathology. The Veterinary Journal 171, 157-165.

Parkin, T.D.H., Clegg, P.D., French, N.P., Proudman, C., Riggs, C.M., Singer, E.R., Webbon, P.M., Morgan, K.L., 2004a. Risk factors for fatal lateral condylar fracture of the third metacarpus/metatarsus in UK racing. Equine Veterinary Journal 37, 192-199.

Parkin, T.D.H., French, N.P., Riggs, C.M., Morgan, K.L., Clegg, P.D., Proudman, C.J., Singer, E.R., Webbon, P.M., 2004b. Risk of fatal distal limb fractures among thoroughbreds involved in the five tpes of racing in the United Kingdom. Veterinary Record 154, 493-497.

Perkins, N.R., Reid, S.W.J., Morris, R.S., 2005a. Risk factors for injury to the superficial digital flexor tendon and suspensory apparatus in Thoroughbred racehorses in New Zealand. New Zealand Veterinary Journal 53, 184-192.

Perkins, N.R., Reid, S.W.J., Morris, R.S., 2005b. Risk factors for musculoskeletal injuries of the lower limbs in Thoroughbred racehorses in New Zealand. New Zealand Veterinary Journal 53, 171-183.

Philipsson, J., Andreasson, E., Sandgren, B., Dalin, G., Carlsten, J., 1993. Osteochondrosis in the tarsocrural joint and osteochondral fragments in the fetlock joints in Standardbred trotters. II. Heritability. Equine Veterinary Journal 25, 38-41.

Pieramati, C., Pepe, M., Silvestrelli, M., Bolla, A., 2003. Heritability estimation of osteochondrosis dissecans in Maremmano horses. Livestock Production Science 79, 249-255.

Pinchbeck, G., 2004. Horse injuries and racing practices in National Hunt racehorses in the UK: the results of a prospective cohort study. The Veterinary Journal 167, 45-52.

Pritchard, T., Coffey, M., Mrode, R., Wall, E., 2012. Genetic parameters for production, health, fertility and longevity traits in dairy cows. Animal, 1-13.

Ramseyer, A., Gaillard, C., Burger, D., Straub, R., Jost, U., Boog, C., Marti, E., Gerber, V., 2007. Effects of Genetic and Environmental Factors on Chronic Lower Airway Disease in Horses. Journal of Veterinary Internal Medicine, 149-156.

Ramzan, P.H.L., Palmer, L., 2011. Musculoskeletal injuries in Thoroughbred racehorses: A study of three large training yards in Newmarket, UK (2005-2007). Veterinary Journal 187, 325-329.

Ramzan, P.H.L., Powell, S.E., 2010. Clinical and imaging features of suspected prodromal fracture of the proximal phalanx in three Thoroughbred racehorses. Equine Veterinary Journal 42, 164-169.

Reardon, R.J., Boden, L.A., Mellor, D.J., Love, S., Newton, J.R., Stirk, A.J., Parkin, T.D., 2012. Risk factors for superficial digital flexor tendinopathy in Thoroughbred racehorses in hurdle starts in the UK (2001-2009). Equine Vet J 44, 564-569.

Reed, S.R., Jackson, B.F., Mc Ilwraith, C.W., Wright, I.M., Pilsworth, R., Knapp, S., Wood, J.L., Price, J.S., Verheyen, K.L., 2012. Descriptive epidemiology of joint injuries in Thoroughbred racehorses in training. Equine Vet J 44, 13-19.

Reilly, G.C., Currey, J.D., Goodship, A.E., 1997. Exercise of young thoroughbred horses increases impact strength of the third metacarpal bone. Journal of Orthopaedic Research 15, 862-868.

Rieder, S., Hagger, C., Obexer-Ruff, G., Leeb, T., Poncet, P.A., 2008. Genetic Analysis of White Facial and Leg Markings in the Swiss Franches-Montagnes Horse Breed. Journal of Heredity 99, 130-136.

Riggs, C.M., 1999. Aetiopathogenesis of parasagittal fractures of the distal condyles of the third metacarpal and third metatarsal bones - review of the literature. Equine Veterinary Journal 31, 116-120.

Riggs, C.M., 2002. Fractures - A preventable hazard of racing thoroughbreds? Veterinary Journal 163, 19-29.

Riggs, C.M., Whitehouse, G.H., Boyde, A., 1999. Pathology of the distal condyles of the third metacarpal and third metatarsal bones of the horse. Equine Veterinary Journal 31, 140-148.

Santschi, E.M., Purdy, A.K., Valberg, S.J., Vrotsos, P.D., Kaese, H., Mickelson, J.R., 1998. Endothelin receptor B polymorphism associated with lethal white foal syndrome in horses. Mammalian Genome 9, 306-309.

Schroder, W., Klostermann, A., Distl, O., 2011. Candidate genes for physical performance in the horse. Veterinary Journal 190, 39-48.

Schroderus, E., Ojala, M., 2010. Estimates of genetic parameters for conformation measures and scores in Finnhorse and Standardbred foals. J Anim Breed Genet 127, 395-403.

Schroter, R.C., Marlin, D.J., Denny, E., 1998. Exercise-induced pulmonary haemorrhage (EIPH) in horses results from locomotory impact induced trauma--a novel, unifying concept. Equine Vet J 30, 186-192.

Schurink, A., Ducro, B.J., Heuven, H.C.M., van Arendonk, J.A.M., 2011. Genetic parameters of insect bite hypersensitivity in Dutch Friesian broodmares. Journal of Animal Science 89, 1286-1293.

Shakhsi-Niaei, M., Klukowska-Rotzler, J., Drogemuller, C., Swinburne, J.E., Gerber, V., Leeb, T., 2010. Characterization of the equine ITGAX gene and its association with recurrent airway obstruction in European Warmblood horses. Anim Genet 41, 559-560.

Shelton, D.R., Martin, R.B., Stover, S.M., Gibeling, J.C., 2003. Transverse fatigue crack propagation behavior in equine cortical bone. Journal of Materials Science 38, 3501-3508.

Shi, L., Wang, D., Riggs, C.M., Qin, L., Griffith, J.F., 2011. Statistical analysis of bone mineral density using voxel-based morphometry-an application on proximal sesamoid bones in racehorses. Journal of Orthopaedic Research 29, 1230-1236.

Sobczynska, M., 2010. Genetic parameters of racing performance indices in Polish Arabian horses. Livestock Science 131, 245-249.

Stephen, J.O., White, N.A., 2nd, McCormick, W.H., Cowles, R.R., Corley, K.T., 2003. Risk factors and prevalence of injuries in horses during various types of steeplechase races. Journal of the American Veterinary Medical Association 223, 1788-1790.

Stepnik, M.W., Radtke, C.L., Scollay, M.C., Oshel, P.E., Albrecht, R.M., Santschi, E.M., Markel, M.D., Muir, P., 2004. Scanning electron microscopic examination of third metacarpal/third metatarsal bone failure surfaces in thoroughbred racehorses with condylar fracture. Veterinary Surgery 33, 2-10.

Stewart, I.D., White, I.M.S., Gilmour, A.R., Thompson, R., Woolliams, J.A., Brotherstone, S., 2012. Estimating variance components and predicting breeding values for eventing disciplines and grades in sport horses. Animal 6, 1377-1388.

Stewart, I.D., Woolliams, J.A., Brotherstone, S., 2010. Genetic evaluation of horses for performance in dressage competitions in Great Britain. Livestock Science 128, 36-45.

Stock, K.F., Distl, O., 2006. Genetic correlations between osseous fragments in fetlock and hock joints, deforming arthropathy in hock joints and pathologic changes in the navicular bones of Warmblood riding horses. Livestock Science 105, 35-43.

Stock, K.F., Hoeschele, I., Distl, O., 2007. Estimation of genetic parameters and prediction of breeding values for multivariate threshold and continuous data in a simulated horse population using Gibbs sampling and residual maximum likelihood. Journal of Animal Breeding and Genetics 124, 308-319.

Sun, C., Madsen, P., Nielsen, U.S., Zhang, Y., Lund, M.S., Su, G., 2009. Comparison between a sire model and an animal model for genetic evaluation of fertility traits in Danish Holstein population. Journal of Dairy Science 92, 4063-4071.

Suontama, M., van der Werf, J.H.J., Juga, J., Ojala, M., 2012. Genetic parameters for racing records in trotters using linear and generalized linear models. Journal of Animal Science 90, 2921-2930.

Swinburne, J.E., Bogle, H., Klukowska-Rötzler, J., Drögemüller, M., Leeb, T., Temperton, E., Dolf, G., Gerber, V., 2009. A whole-genome scan for recurrent airway obstruction in Warmblood sport horses indicates two positional candidate regions. Mammalian Genome 20, 504-515.

Takahashi, T., Hiraga, A., Ohmura, H., Kai, M., Jones, J.H., 2001. Frequency of and risk factors for epistaxis associated with exercise-induced pulmonary hemorrhage in horses: 251,609 race starts (1992-1997). Journal of the American Veterinary Medical Association 218, 1462-1464.

Thomson, P.C., Wilson, B.J., Wade, C.M., Shariflou, M.R., James, J.W., Tammen, I., Raadsma, H.W., Nicholas, F.W., 2010. The utility of estimated breeding values for inherited disorders of dogs. The Veterinary Journal 183, 243-244.

Thorpe, C.T., Clegg, P.D., Birch, H.L., 2010. A review of tendon injury: why is the equine superficial digital flexor tendon most at risk? Equine Vet J 42, 174-180.

Tozaki, T., Hill, E.W., Hirota, K., Kakoi, H., Gawahara, H., Miyake, T., Sugita, S., Hasegawa, T., Ishida, N., Nakano, Y., Kurosawa, M., 2012. A cohort study of racing performance in Japanese Thoroughbred racehorses using genome information on ECA18. Anim Genet 43, 42-52.

van der Werf, J., 2012. Mixed Models for Genetic Analysis. VSN International Ltd.

van Grevenhof, E.M., Schurink, A., Ducro, B.J., van Weeren, P.R., van Tartwijk, J.M.F.M., Bijma, P., van Arendonk, J.A.M., 2009. Genetic variables of various manifestations of osteochondrosis and their correlations between and within joints in Dutch warmblood horses. Journal of Animal Science 87, 1906-1912.

Vazquez, A.I., Gianola, D., Bates, D., Weigel, K.A., Heringstad, B., 2009. Assessment of Poisson, logit, and linear models for genetic analysis of clinical mastitis in Norwegian Red cows. Journal of Dairy Science 92, 739-748.

Verheyen, K., Newton, J., Price, J., Wood, J., 2006. A case-control study of factors associated with pelvic and tibial stress fractures in Thoroughbred racehorses in training in the UK. Preventive Veterinary Medicine 74, 21-35.

Verheyen, K., Price, J., Wood, J., 2007. Fracture rate in Thoroughbred racehorses is affected by dam age and parity. The Veterinary Journal 174, 295-301.

Verheyen, K.L., Wood, J.L., 2004. Descriptive epidemiology of fractures occurring in British Thoroughbred racehorses in training. Equine Vet J 36, 167-173.

Visscher, P.M., Hill, W.G., Wray, N.R., 2008. Heritability in the genomics era--concepts and misconceptions. Nat Rev Genet 9, 255-266.

Wade, C.M., Giulotto, E., Sigurdsson, S., Zoli, M., Gnerre, S., Imsland, F., Lear, T.L., Adelson, D.L., Bailey, E., Bellone, R.R., Blocker, H., Distl, O., Edgar, R.C., Garber, M., Leeb, T., Mauceli, E., MacLeod, J.N., Penedo, M.C.T., Raison, J.M., Sharpe, T., Vogel, J., Andersson, L., Antczak, D.F., Biagi, T., Binns, M.M., Chowdhary, B.P., Coleman, S.J., Della Valle, G., Fryc, S., Guerin, G., Hasegawa, T., Hill, E.W., Jurka, J., Kiialainen, A., Lindgren, G., Liu, J., Magnani, E., Mickelson, J.R., Murray, J., Nergadze, S.G., Onofrio, R., Pedroni, S., Piras, M.F., Raudsepp, T., Rocchi, M., Roed, K.H., Ryder, O.A., Searle, S., Skow, L., Swinburne, J.E., Syvanen, A.C., Tozaki, T., Valberg, S.J., Vaudin, M., White, J.R., Zody, M.C., Lander, E.S., Lindblad-Toh, K., 2009. Genome Sequence, Comparative Analysis, and Population Genetics of the Domestic Horse. Science 326, 865-867.

Watkins, K., Stewart, B. 2008. EIPH and horseracing in Hong Kong: Scale of the problem, management, regulation and unique aspects. In: Proceedings of a Workshop on exercise-induced pulmonary haemorrhage: state of current knowledge, Granville Island, Vancouver, Canada, 52-56.

Weatherby, J., 1791. An introduction to a general stud-book : containing (with few exceptions) the pedigree of every horse, mare, &c. of note, that has appeared on the turf for the last fifty years, with many of an earlier date; together with a short account of the most noted Arabians, Barbs, &c. connected therewith.

Webbon, P., 2012. Harnessing the genetic toolbox for the benefit of the racing Thoroughbred. Equine Vet J 44, 8-12.

Weideman, H., Schoeman, S.J., Jordaan, G.F., 2004. A genetic analysis of epistaxis as associated with EIPH in the Southern African Thoroughbred. South African Journal of Animal Science 34, 265-273.

West, J.B., Mathieu-Costello, O., 1994. Stress failure of pulmonary capillaries as a mechanism for exercise induced pulmonary haemorrhage in the horse. Equine Vet J 26, 441-447.

Whitton, R.C., Trope, G.D., Ghasem-Zadeh, A., Anderson, G.A., Parkin, T.D., Mackie, E.J., Seeman, E., 2010. Third metacarpal condylar fatigue fractures in equine athletes occur within previously modelled subchondral bone. Bone 47, 826-831.

Wilke, V.L., Conzemius, M.G., Kinghorn, B.P., Macrossan, P.E., Cai, W., Rothschild, M.F., 2006. Inheritance of rupture of the cranial cruciate ligament in Newfoundlands. Journal of the American Veterinary Medical Association 228, 61-64.

Williams, K.J., Derksen, F.J., de Feijter-Rupp, H., Pannirselvam, R.R., Steel, C.M., Robinson, N.E., 2008. Regional pulmonary veno-occlusion: a newly identified lesion of equine exercise-induced pulmonary hemorrhage. Veterinary Pathology 45, 316-326.

Williams, R.B., Harkins, L.S., Hammond, C.J., Wood, J.L., 2001. Racehorse injuries, clinical problems and fatalities recorded on British racecourses from flat racing and National Hunt racing during 1996, 1997 and 1998. Equine Vet J 33, 478-486.

Wilsher, S., Allen, W.R., Wood, J.L.N., 2006. Factors associated with failure of Thoroughbred horses to train and race. Equine Veterinary Journal 38, 113-118.

Wittwer, C., Dierks, C., Hamann, H., Distl, O., 2008. Associations between Candidate Gene Markers at a Quantitative Trait Locus on Equine Chromosome 4 Responsible for Osteochondrosis Dissecans in Fetlock Joints of South German Coldblood Horses. Journal of Heredity 99, 125-129.

Ytrehus, B., Carlson, C.S., Ekman, S., 2007. Etiology and pathogenesis of osteochondrosis. Veterinary Pathology 44, 429-448.