Universidade de Évora - Instituto de Investigação e Formação Avançada

Programa de Doutoramento em Ciências Veterinárias

Tese de Doutoramento

Application of objective and non-invasive methodologies for welfare assessment in endurance horses

Mónica Alexandra Freire Cardoso de Mira

Orientador(es) | Elsa Cristina Carona de Sousa Lamy
Rute Guedes dos Santos

Évora 2021

Universidade de Évora - Instituto de Investigação e Formação Avançada

Programa de Doutoramento em Ciências Veterinárias

Tese de Doutoramento

Application of objective and non-invasive methodologies for welfare assessment in endurance horses

Mónica Alexandra Freire Cardoso de Mira

Orientador(es) | Elsa Cristina Carona de Sousa Lamy
Rute Guedes dos Santos

Évora 2021

A tese de doutoramento foi objeto de apreciação e discussão pública pelo seguinte júri nomeado pelo Diretor do Instituto de Investigação e Formação Avançada:

Presidente	Elisa Maria Bettencourt (Universidade de Évora)
Vogais	Ester Bartolome Medina () Luis Pardon Lamas (Universidade de Lisboa - Faculdade de Medicina Veterinária) Mário Gonçalves Cotovio (Universidade de Trás-os-Montes e Alto Douro) Rute Guedes dos Santos (Escola Superior Agrária de Elvas (ESAE/IPPortalegre)) (Orientador) Susana Monteiro (Universidade de Évora)

Évora 2021

Quanto é Alto e Régio o Pensamento

Ponho na altiva mente o fixo esforço
Da altura, e à sorte deixo,
E as suas leis, o verso;

Que, quanto é alto e régio o pensamento,
Súbita a frase o busca
E o 'scravo ritmo o serve.

Ricardo Reis *in* Odes (1924)

(...) there is no religion without love, and people may talk as much as they like about their religion, but if it does not teach them to be good and kind to man and beast, it is all a sham (...)

Anne Sewell *in* Black Beauty (1877)

ACKNOWLEDGEMENTS

I want to thank my supervisors, Elsa Lamy, David Marlin and Rute Santos, for their shared knowledge, support and constructive criticism. Also, for agreeing to supervise my work in sometimes challenging conditions.

I extend my gratitude to the many colleagues from different countries that have helped me throughout all the fieldwork, data analysis and writing of the papers included in this thesis. A special thank you to Marco Lopes, Carina Santos, Ângela Eleutério, Jane Williams, Patrícia Rodrigues, Beatriz Arroja, Orlando Fernandes, Constanza Gómez-Alvarez, Massimo Puccetti, Mafalda Vaz Pinto, Pedro Martins, Irina Pinheiro and Rita Delgado.

Many thanks to Massimo Pucetti, Emma Persson Sjödin, Marie Rhodin and Filipe Serra Bragança, for their generosity.

I also would like to thank all the organizing committees of endurance competitions, endurance stables and owners, and the Fédération Equestre International to have agreed with some of the studies presented here.

Thank you also to my co-workers who brilliantly supported my company during the countless day-offs, many more than expected, needed to complete this thesis.

Finally, to my friends and family, for being there, unconditionally, for me.

ABSTRACT

Endurance riding evolved in the last two decades from an amateur activity into a highly professionalised sport. Better training techniques and more specialised breeding allowed the creation of super endurance equine athletes, capable of achieving a sustained high speed along with a fast-cardiac recovery capacity. However, the competitiveness and inherent effort of the sport lead to the persistence, up to the present, of severe injuries, unacceptable current societal standards, despite the intense veterinary monitoring of horses in competitions. This research aimed to study non-invasive methods that could be used in an endurance competition setting to objectively measure performance and distress. Salivary cortisol, eye temperature measured by infrared thermography and gait quantification through an inertial sensor-based system were studied to determine the practicality, reliability and repeatability of results in competitions and experimental settings. The objective results were confronted with the subjective evaluations and the outcome of horses. Firstly, the receptiveness by veterinarians to the utilization of objective methods in competitions was assessed. The results showed that two thirds of 157 Federation Equestre Internationale (FEI) official endurance veterinarians would be receptive to a new reliable technology to objectively quantify gait. Indeed, a trial performed during an endurance ride using a sensor-based system had a favourable impression regarding its logistic use, but the objective results indicated a significant disagreement in relation to the subjective gait assessment of veterinarians. To assess whether the delay caused by instrumentation of the horse with the sensors would affect the cardiac recovery index (CRI), an early indicator of fatigue in horses used by veterinarians in competitions, another study was performed. No effect of a waiting time was found on the CRI. A second study using the sensor-based system was performed next, to assess the repeatability and agreement of results of subjective and objective gait assessment with different handlers and trot-up presentation styles. It was demonstrated that veterinarians gave better scores to horses trotted as for a competition when compared to a regular presentation for a lameness work-up. The objective gait results were not affected. Finally, the salivary cortisol and ocular temperature measured by infrared thermography were assessed in competitions to investigate whether there was a relation to the outcome of competing horses. It was shown that these biomarkers could potentially be used in association to characterize physical effort and emotional stress in endurance competitions, but further studies are needed. Globally, the results of this research showed the potential role of non-invasive methods in competitions to better objectify the results and as means to obtain relevant information for the endurance sport guidelines.

Key words: endurance; salivary cortisol; eye temperature; objective gait analysis; lameness; horse;

RESUMO

A resistência equestre (Endurance) evoluiu nas últimas duas décadas de uma atividade amadora para um desporto altamente profissionalizado. A melhoria das técnicas de treino e a criação mais especializada permitiram a criação de superatletas equinos de Endurance, capazes de atingir velocidades elevadas sustentadas em conjunto com uma recuperação cardíaca rápida. Porém, a competitividade e o esforço inerentes ao desporto levam à ocorrência de lesões graves, inaceitáveis perante os padrões sociais atuais, apesar do intenso acompanhamento veterinário dos cavalos durante as competições. Esta investigação teve como objetivo estudar métodos não invasivos que possam ser usados em ambiente de competição de Endurance para medir objetivamente o desempenho e o stresse. O cortisol salivar, a temperatura ocular medida por termografia infravermelha e a quantificação dos andamentos através de um sistema baseado em sensores de inércia foram estudados para determinar a exequibilidade, confiabilidade e repetibilidade dos resultados em competições e ambientes experimentais. Os resultados objetivos foram confrontados com as avaliações subjetivas e os resultados dos cavalos. Primeiramente, avaliou-se a recetividade dos médicos veterinários à utilização de métodos objetivos nas competições. Os resultados mostraram que dois terços de uma amostra de 157 veterinários oficiais de Endurance da Federação Equestre Internacional (FEI) estariam recetivos a uma nova tecnologia fiável para quantificar objetivamente a marcha. Um ensaio realizado durante uma prova de resistência usando um sistema baseado em sensores provocou uma impressão favorável quanto aos aspetos logísticos, mas os resultados objetivos indicaram uma discordância significativa em relação à avaliação subjetiva dos andamentos realizada pelos veterinários. Realizou-se um outro estudo para avaliar se o tempo despendido na instrumentação do cavalo com os sensores afetaria o índice de recuperação cardíaca (IRC), um indicador precoce de fadiga em cavalos utilizado por veterinários em competições. Verificou-se que o IRC não foi influenciado pelo tempo de espera. Em seguida, realizou-se um segundo estudo utilizando o sistema baseado em sensores, para avaliar a repetibilidade e a concordância dos resultados da avaliação subjetiva e objetiva da marcha com diferentes apresentadores e estilos de apresentação a trote. Foi demonstrado que os veterinários pontuaram melhor os cavalos trotados no estilo utilizado em competição quando comparados ao estilo utilizado numa apresentação para um exame de claudicação. Os resultados objetivos da marcha não foram afetados. Finalmente, o cortisol salivar e a temperatura ocular medidos por termografia infravermelha foram avaliados em competições, para avaliar se existia relação com os resultados de cavalos em competição. Ficou demonstrado que estes biomarcadores poderiam ser usados em associação para caracterizar o esforço físico e o stresse emocional em competições de Endurance, mas são necessários estudos adicionais. Globalmente, os resultados desta investigação mostraram o potencial papel dos métodos não invasivos para melhor objetivar os resultados em competição e como meio de obter informações relevantes para as diretrizes do desporto de resistência equestre.

Palavras-chave: resistência equestre; cortisol salivar; temperatura ocular; análise objetiva dos andamentos; claudicação; cavalo.

The doctoral thesis is structured in seven chapters. Chapter I consists of a general introduction that highlights the background and the main aims of the research. This chapter also provides the current framing of endurance as an equestrian sport, including a relevant bibliographic review of the research's subject areas.

Chapters II, III, IV, V and VI correspond to each of the scientific publications generated by this research. Chapters II, III and IV include copies of the published papers, whereas chapters V and VI correspond to the publications submitted and currently under review by the journals. The thesis addressed three main research areas: objective gait analysis, infra-red eye thermography and salivary cortisol.

The gait assessment is addressed in chapters II, III and V. Chapter II includes the published paper "Challenges encountered by Federation Equestre Internationale (FEI) veterinarians in gait evaluations during FEI endurance competitions: an international survey" that investigated in detail the limitations felt by veterinarians while judging gait in endurance competitions. This publication is essential because it provides evidence-based information about the importance of different factors in the subjective gait assessment, whose description had been only reported anecdotally until now. Moreover, it assesses veterinarians' receptiveness to introducing new technologies that allow an objective analysis of the gait.

Chapter III includes the published paper "Objective detection and quantification of irregular gait with a portable inertial sensor-based system in horses during an endurance race – a preliminary assessment", which assessed the feasibility of using a sensor-based portable gait analysis system in an authentic setting, e.g. in the veterinary inspections during two endurance rides. The study identified was able to identify technical limitations of the use of such a system regarding the logistics of a veterinary inspection, receptiveness by riders and generate preliminary comparison results between the subjective and objective results.

Chapter IV contains the published paper "Do waiting times in endurance vet gates affect the cardiac recovery index?". The study explored one of the limitations encountered by the previous chapter study, e.g. the introduction of a necessary time delay due to the horses' instrumentation with sensors, between the two heart rate counts used to calculate the Cardiac Recovery Index (CRI), an important indicator used by veterinarians to assess the metabolic fitness to compete of horses in endurance rides. To this end, we studied how "natural" delays of the 60 seconds preconised to perform the CRI related to logistic difficulties of the veterinary inspections occurring commonly in real competitions (e.g. too many horses for the available veterinarians) affected the heart rate counts. Moreover, the publication provided important conclusions regarding the basal heart rate and its relationship to the

CRI result and interpretation. It also brought to the Fédération Equestre Internationale (FEI) attention that the CRI guidelines to veterinarians were not accurate, subsequently changed in the rules.

Chapter V includes the submitted paper "Vet gate trotting style improves subjective gait grading in endurance horses compared to a lameness presentation style". This resulted from an experimental study conducted to address the most important finding of the study presented in Chapter II, which identified the horse's handling as the main factor hindering a proper gait assessment by veterinarians in competitions. Moreover, this study investigated the veterinarian's agreement in gait assessment under different conditions and using different grading systems providing further guidance for the definition of the ABC score system used in endurance competitions.

Finally, chapter VI presents the paper "Salivary cortisol and infrared thermographic ocular temperature use as biomarkers during endurance competitions", which investigated the usefulness of two different non-invasive biomarkers to assess competing horses' fitness and welfare, namely salivary cortisol and infra-red eye temperature.

To conclude, Chapter VII presents the research's most significant conclusions, highlighting its practical applications and future research directions to safeguard endurance horses' welfare in competition.

LIST OF FIGURES

CHAPTER VI: Salivary Cortisol and Infrared Thermographic Ocular Temperature Use as Biomarkers During Endurance Competitions

LIST OF TABLES

CHAPTER III: Objective Detection and Quantification of Irregular Gait with a Portable Inertial Sensor-Based System in Horses during an Endurance Race – A Preliminary Assessment

CHAPTER IV: Do Waiting Times in Endurance Vet Gates Affect the Cardiac Recovery Index?

CHAPTER V: Vet gate trotting style improves subjective gait grading in endurance horses when compared to a lameness presentation style

CHAPTER VI: Salivary Cortisol and Infrared Thermographic Ocular Temperature Use as Biomarkers During Endurance Competitions

CHAPTER I: General Introduction

1. Overview on Endurance Sport and Equine Welfare

Endurance rides are long-distance races designed to test the speed and stamina of a horse, in addition to the rider's capacity to conduct a horse across all kinds of terrain under various meteorological conditions. Compared to human marathons or trail races, equestrian rides are not continuous races but divided into phases followed by a compulsorily rest. A veterinary inspection is mandatory before the ride and after each phase, to ensure competing horses' welfare. The word endurance means the ability to bear suffering. Adopted worldwide to design long-distance competitions, it seems to have its etymology in old French ‘durance’ (duration), descending from the Latin verb *durare* or making hard. In turn, the words ride, and race seem to have old English and Germanic roots (https://www.etymonline.com). Ride (sitting on a horse) was adopted worldwide and adapted to the phonetics of Latin originated languages and popularized as “raid(e)” to designate an endurance competition. In English, endurance ride and race (contest of speed) are used interchangeably by the endurance community, being the first connected to a more conservative approach to the sport and the second with the faster flat tracks of the desert endurance in the Middle East.

Equine welfare dates back far as the early 19th century, with the Society for the Prevention of Cruelty to Animals (SPCA) funded in 1824 (SPCAI, 2021). Since the early days, endurance rides were associated with horse’s fatalities generating public hostility. Protection of horses used in sport and racing is an incrementally topical subject in today’s society. Although being by far the discipline most closely monitored by veterinarians, its increasing competitivity has been precluding the desired results crucial for its survival in the modern world, e.g., reducing competing horses’ morbidity and absence of fatalities. Evidence-based regulation changes to protect the horse have been the scope in the last decade of the Fédération Equestre Internationale (FEI), the international regulatory body of equine sports. The introduction of objective non-invasive methodologies to quantify the response to competing horses’ exercise could, furthermore, provide unbiased information to veterinarians working in competitions, being this thesis’ research working scope.

2. Historical Perspective of Endurance Riding

2.1 The Early Days

Horses were once the main mean of man transportation and communication, with George Washington reporting associated metabolic disorders already in the 18th century. Allegedly his horse foundered badly after a 128 km ride journey, and another horse died of a heat stroke (Frazier, 2000).

Recreational endurance rides were interestingly reported in several countries worldwide even before motorization, and the horse had lost its utility as a workforce to humankind. As expectable, the cavalry was in the origins of competitive endurance rides in Europe. The aim was mainly to test endurance in horses, an important feature for the cavalry. Competitive rides between countries were also organized. One of the most covered endurance rides was the *Distanzritt* between Berlin and Vienna, comprising 630 Km with an average distance of 120 km per day for five days. In 1892, a death toll of 25 horses, out of 200 starters and 145 finishers, dying in the track was reported (Jurga, 2016). The horses, mostly Thoroughbreds and some Hungarian Arabian Shagyas, were given morphine to keep going, and the winner horse died a few hours later after passing the finish line (Crockett, 2018). Meanwhile, to prove western cow ponies were most suitable for the cavalry Buffalo Bill, a well-known figure of the American Old West, organized in 1893 the American Great Cowboy Race. This was a civilian 1000-mile ride from Nebraska to Chicago, the host of the World Fair in this year. Already by then, this brought great concern to the animal rights associations, who tried to stop the race considering the recent events in the *Distanzritt* in Europe (Serrano, 2016). For the Great Race to take place, the organizing commission was forced to negotiate control points along the route, where humane society veterinarians would examine the horses to verify there was no threat to their health. The first veterinary inspections in the history of endurance were born, and for the first time, horses could be disqualified if found to be unsound or show signs of abuse (Bache, 2017). Furthermore, humane society officials accompanied each of the nine competitors riding in a buggy in the last part of the route, and horses were examined after the ride (Serrano, 2016).

Endurance became increasingly popular in the first half of the XX century with ordinary citizens taking progressively over the rides. Most rides took place in a continuous track between cities (Brussels-Oostende, Madrid-Lisbon, Budapest-Vienna, Paris-Deauville) without defined distance or rules. In Uruguay, one of the countries with long-standing tradition in endurance riding, long-distance races were born from ordinary people, most likely inspired by the *caballerias gauchas*. The first long distance competition of 80 km, locally named as *raid*, was reported in 1913 with a speed of 48 km/h to become a national sport in 1935, dragging crowds and broadcasted live on the radio and TV until today (Maisonnave and Lockhart, 2012). The *raid* still exists today, but it has different rules from today's endurance riding. For example, continued crewing is allowed with cooling being performed from hoses from the trucks that closely follow the horses instead of crewing points. Most likely, the *raid* represented a major influence on today's desert endurance rides, including the desert style of riding, meanwhile also widespread worldwide, with the rider sitting back in the saddle with the legs pulled forward on long stirrups.

2.2 Birth of Endurance as a Sport

Perhaps the most instrumental rule to prevent horses' overexertion used until present times was introducing an ingenious gate into hold after each phase of a ride. Therewith horses had to meet a pulse criterion before the competition time stopped at the end of a phase, e.g., the time spent to get the pulse down to a pre-established level was added to the time on the track. After the finish line of a phase, competitors passed another timing gate. Veterinarians measured the heart rate and, once the pulse criteria were met, the veterinary inspection was a brief physical examination that included a trot-up to assess gait. This area has been named the vet gate until present times. The maximum heart rate allowed for a horse to recover and attain it to be kept in the competition was first 72 bpm over 45 minutes in 1962; it was then reduced to 68 bpm with shorter recovery times (Nicholson, 2007), and finally established at 64 bpm within 30 minutes. A vet card was also introduced to register the clinical parameters numerically or coded into A, B and C, and passed along to the next veterinarian for consultation in the next vet gate. Kerry Ridgeway, together with other founding veterinary members of the AERC (Mackay-Smith M., 2016), was immortalized by the Cardiac Recovery Index (CRI), also known as the Ridgeway Trot, a parameter that consists of two heart rate counts separated by one minute time during each the horse was trotted up. Early observations showed that a fit horse would keep a pulse at the second count not higher than 4bpm than the first count (Ridgeway, 1988). Other innovative rule changes to protect horses from overexertion introduced a progressive distance mandatory qualifying system that partially avoided unprepared horses to participate in long rides and the introduction of the minimum age of five years to compete.

The concept gained worldwide acceptance and soon was exported to other regions such as Europe, Australia and South Africa, all founding their associations and iconic rides that subsisted until today (the 160 km Tom Quilty Cup in Australia and the 3-day x 80 km Fauresmith Endurance Ride in South Africa) with their rules being widely inspired in the AERC. In Europe, the first modern rides arise in France, with some of those competitions also subsisting until today (the 160 km of Florac and the 2-day 100 km Montcuq endurance rides). The Endurance and Long-Distance Riding International Conference (ELDRIC), an association strongly inspired by the AERC, emerged in Europe to serve as an umbrella for European associations and competitions (Ancelet, 1986). Though a short-lived but highly dynamic endurance association (1979-2003), the ELDRIC was pivotal for developing endurance in Europe. It held an annual conference that encouraged research, generated scientific publications based on epidemiologic data (Burger and Dollinger, 1988), promoted the education of veterinarians and riders and defined rules for endurance riding in Europe.

As endurance joined the Fédération Equestre Internationale (FEI) in 1982, whose headquarters are based in Europe, the ELDRIC became redundant and eventually faded. The first FEI European and World Championships were held in France and Italy in 1982 and 1986 (FEI, 2020). The FEI rules were, and still are, widely inspired by the ELDRIC/AERC rules.

Around the nineties, as the veterinary inspections became an increasingly critical point to ensure horses' welfare, the competitions started to be progressively organized around the vet gate with ride phases or loops organized as a trefoil, e.g. with horses starting and arriving at the vet gate. Undeniably competitions lost much of their charm with more monotonous tracks, yet this type of set-up soon became the norm, as it was cost-wise more sustainable for organizing committees that did not need to duplicate the veterinary, judges, and timing teams continuous tracks with subsequent vet gates.

2.3 Professionalization of Endurance and Desert Racing

Despite becoming an FEI discipline and increasing popularity, endurance remained a backyard amateur activity for decades. The interest and entrance of the Middle East Countries in the discipline, especially the United Arab Emirates (UAE), changed the discipline paradigm from the mid-nineties until present times. Endurance rides, called then marathons, were first introduced in the early nineties in Qatar, the UAE and later in Bahrain, these three countries being still the most representative countries for the discipline in the Middle East. The enthusiasm was probably linked to the identification of endurance with the Bedouin nomad past, the use of the Arab horse as the most suitable for long-distance riding, and the economic bust of the Middle East countries.

The first FEI World championships in Arabic countries took place in Qatar in 1997 with a 100 Km ride, followed by Abu Dhabi in the United Arab Emirates (UAE) in 1998 in the usual 160 Km format and the World Championships for Juniors and Young Riders in Bahrain in 2005 (Burger and Dollinger, 1988). The UAE established the sport of endurance together with the founding of the Emirati Federation in 1992, the same year the Dubai Racing Club for flat racing was created. Particularly the Emirate of Dubai propelled endurance, with His Highness Sheikh Mohammed bin Rashid Al Maktoum, Vice President and Prime Minister of the UAE and Ruler of Dubai, himself riding in competitions, becoming a matter of national pride. Many royal families in the Middle East embraced the sport, founding their training stables competing between them. Private wealthy owners followed. Endurance stables were built and were mainly inspired by racing stables models. Each stable hosted a team, each one having its brand and uniform. Local trainers from the racing industry or long-standing countries in endurance were specifically hired for each stable. Like in racing, trainers dictated the schedule, and riders became pilots. Many riders and trainers were hired from Europe (France, Spain, Italy, Portugal and later the Eastern European countries such as Hungary, Poland, Bulgaria, Slovakia and the Czech Republic), Latin America (Uruguay, Argentina), Australia, the USA and South Africa. They would then build similar training facilities back home, powered both by horses' sales and training. Breeding for endurance also fired worldwide.

The romantic and more conservative view of endurance riding promoted by the AERC "to finish is to win" and "one rider, one horse" clashed with the new professional endurance-oriented for speed, results and the market. Many amateurs became professionals living from endurance worldwide, entirely dependent on the sales to the Middle East. Training techniques, nutritional and veterinary monitoring allowed endurance horses to reach speeds and performances thought to be physiologically unattainable before. Endurance became the second major discipline after show jumping, overtaking dressage and eventing, and becoming the fastest growing equestrian sport with more riders and horses in the first decade of the 21st century. Lovers of speed, the Middle East constructed specific flat tracks that allowed horses the galop the whole ride. To meet the new endurance, organizing committees started to organize flatter competition in flat terrain to attract Middle East sponsors and riders in their competitions. Only a few historic endurance rides survived the trend. Saddles became very light and adapted to sitting galop. The desert-style riding, e.g., a galop where the rider sits back in the saddle, brings his legs forward to the shoulder blade level with very long stirrups, considered contrary to any classic equitation technique, was adopted worldwide and proved to be efficient (Viry, 2014). Casual wear with training pants and tennis shoes became the mainstream in endurance. Permanent structures only for endurance competitions with state-of-the-art vet gates and electronic timing systems such as the Dubai International Endurance City and the Emirates Endurance Village in Abu Dhabi were built. However, being primarily occupied by deserts and a climate unfavourable to horse breeding, the Middle East had to rely upon horses bred elsewhere.

All this generated an unprecedented market previously inexistent for the endurance horse that boosted the endurance industry, especially in countries with a tradition in breeding, such as France, Australia, South Africa, Uruguay and Argentina. Furthermore, because of the harsh climate from April to October, which precluded competitions during these months, training centres were built in Europe and other countries or horses were given to local training centres to be trained. This also allowed wealthy owners to compete in Europe in Summer. A permanent venue was built in Newmarket in the United Kingdom. The Euston Park allowed the UAE to continue to organise competitions during the summer. Many competitions started to be sponsored by the UAE, Bahrain, Qatar and Oman with prize money, only existent in the Middle East.

All factors quickly led to the professionalisation of the sport in countries with a long-standing tradition in endurance. Inevitably, it also led to a professionals' dependency from the Middle East countries, particularly Dubai, the leading purchaser of endurance horses. The exponential rise of starters and events in the Middle East led to competitions in the endurance countries to qualify and value horses to supply the market. If unawareness of training methods were once the cause of many accidents, primarily metabolic, nowadays, accidents became a consequence of the highly specialised training methods that allowed performances that the musculoskeletal system does not hold and, unfortunately, many times cannot be anticipated. The highly specialised training allowed horses to have unforeseen performances, but musculoskeletal breakdowns such as racing started to occur.

3. Outline of an FEI Endurance Competition

Endurance rides are not continuous races. The competitions are divided into loops ranging from 16 to 40 km each, followed by a veterinary inspection in a designated area known as a vet gate (VG). Once horses cross the finish line of each loop, a limited time of 15 minutes is allowed to meet the pulse criteria, currently 64 bpm. Phase or competition time is therefore constituted by the time spent racing in the track and recovery time, e.g. the time that mediates from crossing the finish line of the phase and entering the vet gate area where the veterinary inspection occurs.

The heart rate is the first parameter to be measured by veterinarians using either an electronic heart rate monitor or a stethoscope. If the pulse criterion is not met, riders have a second chance to present the horse within 15 minutes. A short recovery time is a crucial parameter for success in endurance rides and accounts significantly for the market value of a horse. Therefore, great effort is placed by competitors and crews in getting the pulse down as fast as possible by watering the horses with ice-cold water in the designated recovery area, especially in warm weather. A horse that will not meet pulse criteria fails to qualify for metabolic reasons (FTQ-ME) and needs to be mandatorily examined by a treating veterinarian to decide whether a treatment is needed. If passing successfully the veterinary inspection, a compulsorily rest period between 20 and 60 minutes must occur after that and before the start into another loop (FEI, 2021a).

The last vet gate does not work as a gate into a hold to keep the competitiveness of the last loop, e.g. it is the first dyad rider-horse crossing the finish line is the winner. Nonetheless, a maximum time of 20 minutes to present a horse within pulse criteria is compulsory to validate the qualification and position in the composition.

4. The Veterinary Inspection

Veterinary inspections are compulsory and carried out by a veterinary commission (VC) in the vet gate (VG) to determine if horses are fit to compete or if, after the last phase, they are sound enough to earn the classification obtained when crossing the finish line (FEI, 2021b). The veterinary inspections take place before the competition (preinspection), after each phase (inspection) and 15 minutes before a start into the next phase (re-inspection). In international rides, veterinarians must be accredited by the FEI in Endurance Official Veterinarians (EOVs). Since 2006, veterinarians are classed in a star-system level (2, 3 or 4-star) according to their experience and mandatory training (FEI, 2021b).

Horses can be eliminated by EOVs at any time, even after crossing the finish line, if deemed not to be fit enough to compete. Veterinary inspections assess the metabolic condition of a horse essentially through a physical examination. The first parameter to be measured is the heart rate. Once the pulse criterion, e.g. ≤ 64 bpm, is met, the veterinarian proceeds with the examination to assess the metabolic condition and the soundness of the gait. If the pulse criteria are not met, the horse will immediately exit the vet gate and is allowed to re-present once more (heart rate re-inspection) within the recovery time allowance. The metabolic condition is evaluated by evaluating mucous membranes and capillary refill time, skin tent, and gut sounds. The horse is trotted forth and back in a 40m lane to assess the gait. One minute after the trot-up start, a second rate is taken and subtracted from the first to calculate the cardiac recovery index (CRI). The CRI is not an eliminatory parameter per se but is used as an early indicator of fatigue if higher than 4bpm (Ridgeway, 1991) provided the basal heart rate is ≥ 60bpm (Robert *et al.*, 2002). Caution must be taken not to over-interpret the CRI when the basal heart rate is close to physiologic levels, such as in the re-inspections, where a higher CRI is expected (de Mira, Williams et al. 2020).

A decision to fail a horse has to be legitimated by a panel composed of at least three veterinarians voting anonymously, called upon the request of one veterinarian. Horses are failed to qualify by the VC either for metabolic reasons (FTQ-ME), gait irregularities (FTQ-GA), or an injury such as a traumatic or tack-related wound or soreness (FTQ-MI) and the cause published with the results. Most FTQ-ME are due to high heart rates not meeting the pulse criteria within the time allowance. However, horses might be failed by the veterinary panel even if showing combined signs of fatigue and metabolic morbidity, such as congestive mucous membranes with a capillary refill time n≥ 2 seconds (s), decreased gut sounds, hyperventilation, demeanour and refusal to trot up, and a high CRI.

4.1 Record Format of Clinical Parameters in Endurance

The clinical parameters obtained at each veterinary inspection are registered in a paper-based or electronic veterinary card (vet card) that will follow a horse throughout a competition and be consulted anytime by competitors and veterinarians. The vet card is essential because veterinarians can recall or consult the evolution from the previous vet gate when performed by a colleague. Parameters that can be quantified in heart rate counts, capillary refill time, and skin tent are noted in numbers (beats per minute or seconds). All other qualitative parameters, such as mucous membranes, gut sounds, muscular tonicity and gait, are registered in an A, B, and C score system, originally introduced by Americans in the AERC rules. Briefly, an A score stands for soundness, a B score for an acceptable abnormality to compete and a C score as a marked abnormality. Only the first heart rate count and gait are eliminatory per se. A horse that scores a C in gait will be automatically classed as an FTQ-GA, was necessarily trotted in front of a panel composed by three veterinarians and deemed to anonymously fail, using voting slips (offering options pass/fail), by at least two members. Before voting, the panel can ask for a re-trot, either because the horse was not trotted properly or because one or more members could not make up their minds.

4.2 Interpretation and Subjectivity of Endurance Scores

Currently, written guidelines on how to classify the different parameters are not provided by the FEI in the rules and only briefly referred to in presentations during the two-day mandatory courses for endurance officials held every four years. Most veterinarians will learn from other experienced officials at competitions. However, in countries with solid rider associations with long-standing endurance rules for national competitions, such as the USA and Australia, the meaning of each score is detailed and illustrated with online pictures and videos (AERA, 2020a; AERC, 2016). Only recently, the interpretation of FEI vet gate parameters was published as supplementary material in a peer-reviewed publication (Bennet *et al.*, 2020) without open access and thus not accessible to everyone (Annex II). This might contribute to subjectivity, in particular, gait assessment due to its eliminatory nature and depreciative impact on horses' published performance reports, which often is a cause of dispute of competitors with veterinarians on-site ads on social media (de Mira M.C. *et al.*, 2019).

5. Epidemiological Studies on Elimination Rates in Endurance

Completion rates (starters subtracted from horses failing to complete) in endurance competitions are reported to be traditionally around 50-60% (E. D. Bennet & Parkin, 2018; Burger & Dollinger, 1988; Nagy et al., 2014; Younes et al., 2016) Those rates seem not to have worsened overtime, being mainly dependent on the region where competitions take place (Nagy *et al.*, 2014a) and the inclusion criteria of studies, e.g. whether qualifying rides with controlled speed and lower mileage are included or not in the studies (Table 1). Countries with a long-standing tradition in endurance that favour slow hilly and scenic tracks such as the USA, Australia/New Zealand, and South Africa/Namibia, show higher completion rates than Middle East countries with fast competitive flat tracks. Also Europe with tendentially flatter tracks to meet the horses' sale market, which favour speed, show lower completion rates.

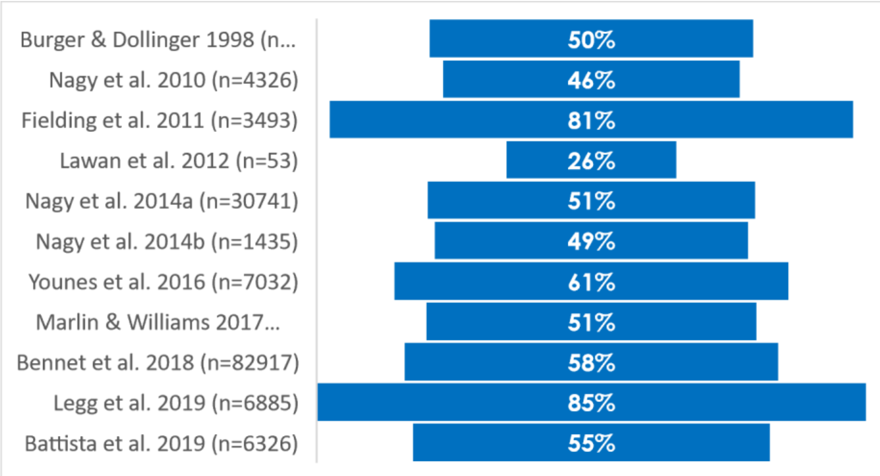


Fig. 1: Reported completion rates in endurance competitions

Most horses (80%) that fail to complete a competition will be eliminated or failed to qualify by veterinarians during the vet gate inspections (Table 2). Endurance has abnormally high elimination rates when compared to other disciplines (Bennet 2017), resulting from the duration and effort of the sport and the number of veterinary controls to avoid severe injuries at all costs. Gait irregularities are the foremost cause of a horse failing to qualify, accounting for two thirds to four fifths of all eliminations (Bennet and Parkin, 2018b; Burger and Dollinger, 1988; Di Battista *et al.*, 2019; Fielding *et al.*, 2011; Marlin and Williams, 2018; Nagy *et al.*, 2013; Nagy *et al.*, 2010; Nagy *et al.*, 2014a; Younes *et al.*, 2016) (Table 3). On the competitor's side, a failure to qualify affects the outcome of a competition, eventually an award or prize money, and can have a significant impact on the horse's market value, causing therefore often resentment towards veterinarian's decisions (Mira *et al.*, 2019).

Endurance riding has since its origins been associated with severe injuries due to the intensity of the effort (Serrano, 2016) However, there has been a shift in the last decades from metabolic injuries with fatal outcomes to stress fractures, similarly to racehorses (Misheff *et al.*, 2010). The discipline evolved in the last two decades from an amateur activity into a highly professionalised sport. The increasing competitiveness of the sport, partly driven by a continuous necessity to promote horses in a discipline traditionally without prize money and, thus, whose professional's livelihood depends on the horse market and training for wealthy owners that want results, still leads, despite stricter and more penalising FEI policies, to the occurrence of musculoskeletal accidents. The current expertise of most professional riders/trainers enables them to recognise a metabolically ill horse in most cases and, therefore, to withdraw voluntarily or accept a decision to eliminate a horse from the competition by a veterinarian. In addition, the perception of a lesser risk for the horse (according to the old saying, "Lame horses don't die") makes the triad rider/trainer/owner much less prone to accept elimination and often elicits complaints towards veterinarians (de Mira M.C. *et al.*, 2019). Moreover, catastrophic musculoskeletal injuries, particularly stress fractures due to a reported mismatch between clinical signs and pending severity, can be hard to predict and, therefore, to prevent (Davidson and Ross, 2003; Milgrom *et al.*, 1985). How many veterinary inspections prevent severe and catastrophic injuries and how many unjustified eliminations occur were not yet quantified and warrant investigation.

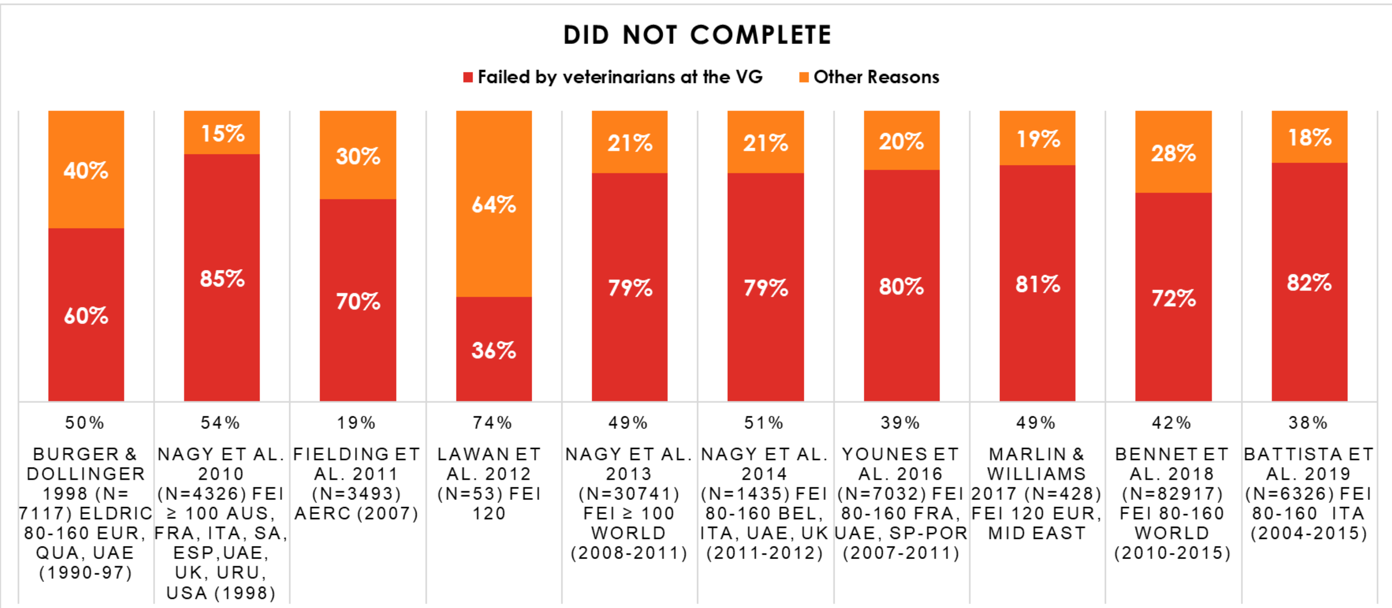


Fig. 2: Published reasons for not having completed a competition - proportion of horses failed at a veterinary vet gate versus other reason

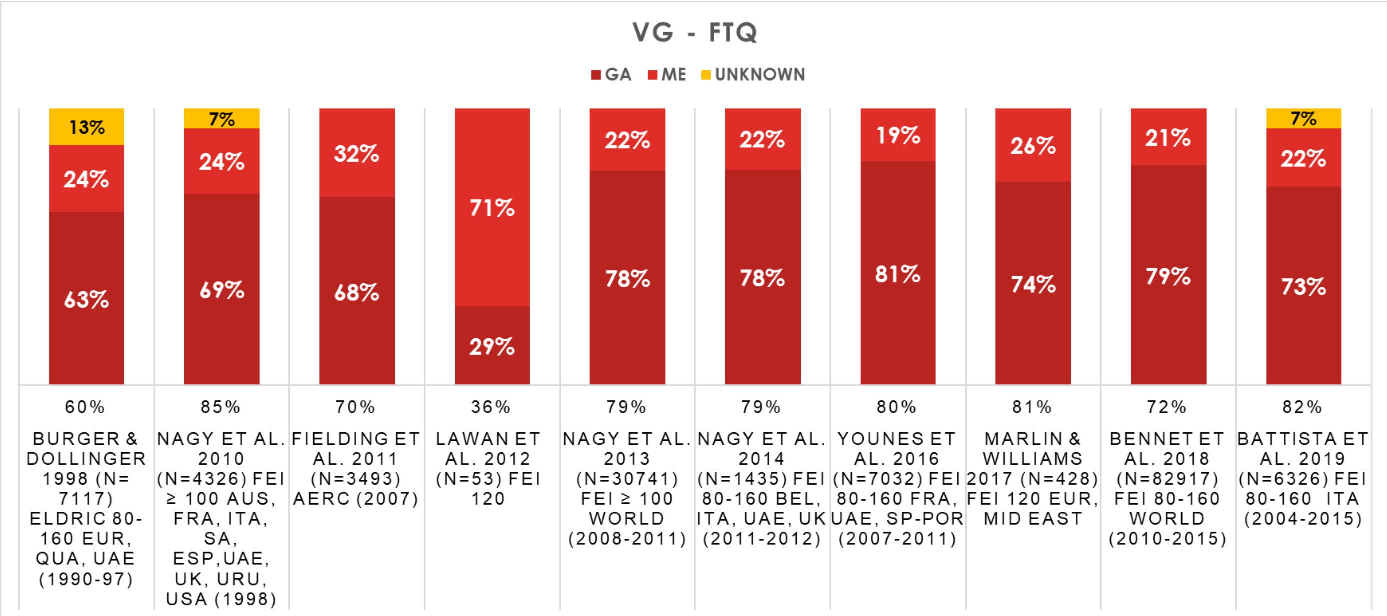


Fig. 3: Published reasons for having failed at a vet gate - gait, metabolic and unknown reasons

6. Factors Identified as Predictive for Non-Completion at Endurance Competitions

Horses will not complete a competition for various reasons, being the most important cause to be failed by veterinarians in inspections at vet gates. A much smaller percentage will not complete because the competitor did not comply with the competition outline, such as being out of time or missing the right track. Various studies investigated which factors related to competition performance (speed and recovery time), vet gate parameters, horse, rider, venue, competition format, the rider would have a predictive value on the horses' outcome (Bennet *et al.*, 2020; Bennet and Parkin, 2018a; b; 2020; Di Battista *et al.*, 2019; Fielding *et al.*, 2011; Nagy *et al.*, 2014a; b; Younes *et al.*, 2015).

The most significant finding was the recent demonstration of the predictive value of a recovery time, the threshold of failing the next vet gate (Younes *et al.*, 2015). This study performed in 7032 starters showed that recovery times of more than 11 and 13 minutes at vet gates 1 or 2 and 3 and 4 would predict in 70% of the cases an FTQ in the next vet gate. Although, as shown previously, speed increases the risk to FTQ (Bennet and Parkin, 2018a; Younes *et al.*, 2016), recovery times predictive value is potentiated but independent of speed (Younes *et al.*, 2015). These studies triggered two recent changes in FEI rules: the reduction of the recovery time allowances and extra penalisations for riders eliminated at more than 20km/h. Other identified factors predictive of non-completion were for horses not having Arabian blood, having participated in a competition in the last 90 days and a previous FTQ history. Previous FTQs and being a male horse or rider also increased the risk (Bennet and Parkin, 2018b; Nagy *et al.*, 2014b). Moreover, a larger field size (> 61 starters), a segment of the track with deep sand, a longest ride class seem to reduce the chances to complete a competition (Bennet and Parkin, 2018b; Fielding *et al.*, 2011; Nagy *et al.*, 2014b; Younes *et al.*, 2016).

7. Morbidity and Fatality Rates at Endurance Competitions

The FEI annual report registered a peak of 20 fatalities from an approximate number of 14 250 starters worldwide in 2012 (0.14%), which steadily decreased to 7 in 16 000 starters in 2019 (0.04%) (FEI, 2019) (Annex 2 a.). A study of 252 738 starts (AERC) in the USA between 2002–2013 showed 67 fatalities in endurance rides, e.g. a fatality rate of 0.03% over 12 years (Balch *et al.*, 2014) , lower than the fatality rate of 0.08% in FEI rides over nine years with 101 fatalities. However, the latter only reflects international but not national competitions, which on the one hand could be lower taking into account the slower and shorter distance rides, but on the other hand, could also be higher due to the faster national races in the Middle East. Interestingly, 45 to 67% of the FEI fatalities were of musculoskeletal origin (Annex 2 b.), while 81% of the AERC fatalities were related to acute abdominal pain with gastric rupture in 12% (Balch *et al.*, 2014). This most likely represents the evolution of endurance rides into much faster races, which brought injuries similar to racehorses resultant from the continuous load, such as stress fractures with similar incidences (Misheff *et al.*, 2010).

Although the exhaustive study of predictive factors of non-completion in endurance rides, an investigation addressing risk factors of morbidity (horses deemed to need a treatment) or fatalities (sudden death or humane euthanasia) in competitions is still missing. Even if far from the fatality rate of 13% (25 of 200 starters) reported at the Vienna-Berlin endurance ride (Jurga 2016) at the end of the 19th century (Serrano, 2016), the current contribution of elimination rates to prevent morbidities or fatalities needs yet to be quantified. Indeed, 46 of the 67 AERC fatalities were horses that were failed by veterinarians, but 21 completed the ride and were judged fit to continue.

8. Subjectivity of Gait Evaluation and Objective Gait Quantification

Until recently, objective gait analysis technology for lameness detection outside the laboratory was impractical and cost-prohibitive because it required expensive equipment and software, complicated horse instrumentation and extensive data collection and analysis (Keegan, 2011; Riber *et al.*, 2006b). New technologies have changed this paradigm, making objective lameness detection and quantification in horses exercising outdoors relatively quick, technically undemanding and inexpensive (Keegan, 2011).

When considering the importance of lameness detection during endurance competitions concerning horse welfare and race outcome, the limitations of lameness detection based exclusively on a brief subjective examination would suggest that using an objective gait analysis system to support lameness detection would likely benefit all segments of the equine endurance industry. A recent study involving FEI veterinarians evaluated the intra/inter-observer agreement of 'fit-to-compete' versus 'non-fit-to compete' judgements and compared the results with a quantitative-gait-analysis system (Sloet Van Oldruitenborgh-Oosterbaan, 2018). The judgment of mild lameness proved difficult between observers, but surprisingly, between observations performed by the same observer over time, although the first evaluation was performed live and the second through video. In fact, perhaps one of the most common complaints following an elimination for lameness is related to the perception of the competitors that the lameness that originated the elimination was the same as in a previous veterinary inspection, where the horse passed, or less than horses perceived as lame (MM, personal observations). Therefore, since a cut-off value in a competition context using objective gait analysis may be hard to establish, objective evolution of lameness during the competition of one horse could be used for veterinarian's consultation. A recent study investigating normal variation between trial, day and horse in gait quantification showed that inter-measurement variation should be expected, even for sound horses. However, less variation was seen within the same horse and increasing repetitions than between different horses (Hardeman, 2018).

9. Biomarkers Used to Assess Welfare in Sport Horses Non-Invasively

Although endurance competitions record the highest elimination rates of all equestrian disciplines and the strictest FEI rules, the recurrence of catastrophic injuries in endurance, particularly musculoskeletal (Bennet and Parkin, 2018b; Marlin and Williams, 2018; Nagy *et al.*, 2014a), frustrates not only competitors but also veterinarians. Moreover, the ongoing social license debate centred on the health and growing welfare concerns with equine athletes arising from the public and society (Heleski *et al.*, 2020; Williams and Marlin, 2020), largely reflected on social media (Campbell, 2016), jeopardise not only horseracing but equestrian sport in general. As a result, there is a current quest for solutions to objectively quantify stress in horses during exercise.

A biomarker can be defined as a characteristic, substance, or process that can objectively be measured and evaluated as an indicator of normal biologic or pathogenic processes and predict the outcome (Strimbu and Tavel, 2010). The attractiveness of their utilisation in sports consists of finding a biomarker that provides an accurate measurement about the compliance of an athlete to the undertaken exercise. This is particularly important in equine athletes because they cannot vocalize distress or pain as humans and cannot decide for themselves (van Loon and Van Dierendonck, 2018). However, biomarker testing poses some challenges in exercise physiology, i.e. limited sensitivity and specificity of single biomarkers to detect injury risk, interindividual variance in absolute values and relative changes; results/reliability are also dependent upon the context and type of exercise, which can result in poorly defined reference ranges for athletes. (Lee *et al.*, 2017). Moreover, exercise is naturally a stressor and induces a biological response to exercise that is difficult to interpret as an enhancer or a limiting factor for the sporting ability of an athlete (Bartolomé and Cockram, 2016).

Furthermore, during competition horses face a mixture of stressors, including transportation (Schmidt *et al.*, 2010b), a new and a noisy environment (Peeters *et al.*, 2013), separation from stable mates (Hartmann *et al.*, 2011) and, specifically in endurance, exposure to large conglomerations of unfamiliar horses in large starts. This complicates the interpretation of biomarker levels because it is hard to differentiate the impact of the different stressors on the welfare from the horses' performance. Musculoskeletal pain from an injury might also arise during a competition (Dyson *et al.*, 2018). However, the impact of pain and discomfort caused by lameness, back-pain, ill-fitted tack and rider on biomarkers is not known in the exercising horse (König v. Borstel *et al.*, 2017).

9.1 Salivary Cortisol

Exhaustedly studied in horses to determine stress levels and the response to different types, intensities and durations of exercise in sport and racehorses, including endurance, cortisol is the end result of the activation of the hypothalamic-pituitary-adrenal (HPA) axis as a response to any psychological or physical stressor. This response is greatly influenced by intrinsic factors (age, gender, breed, inherited temperament, experience) and extrinsic environmental factors. The first studies were performed in plasma, but the identification of the free circulating, i.e. the truly biologically active component of blood cortisol in saliva, and its validation in horses (Peeters *et al.*, 2011), made the collection of this biologic fluid, especially due to its non-invasiveness, much more popular (Peeters *et al.*, 2011). Salivary cortisol (SC) was investigated at rest as a pain-induced marker in the saliva of healthy and diseased horses (Contreras-Aguilar et al., 2019), and also humans (Symons et al., 2015) when vocalization is impaired. Physical activity can dramatically raise the concentration of SC in athletes subjected to different forms of exercise (Nunes *et al.*, 2011). The stress response to different types, intensities and durations of exercise in sport horses was described before (Becker-Birck *et al.*, 2013; Cayado *et al.*, 2006; Janczarek *et al.*, 2013; Jastrzębska *et al.*, 2017; Munk *et al.*, 2017; Peeters *et al.*, 2013; von Lewinski *et al.*, 2013).

It was previously reported that the degree of increase in cortisol seems to reflect better the duration of workload rather than work intensity (Hyyppä, 2005). This was inferred because endurance riding was reported to induce higher increases in cortisol than other equestrian activities such as show jumping, eventing and racing (Desmecht *et al.*, 1996). However, most studies indicate that both high intensity and endurance exercise cause an increase in cortisol in humans (de Graaf-Roelfsema *et al.*, 2007). In endurance exercise, the highest SC increases were reported to occur in the first half of competitions and to stabilise in lower levels (Janczarek *et al.*, 2013; Kędzierski and Cywińska, 2014; Rose *et al.*, 1983). This observation was also reported in human athletes, whose cortisol levels increased after short-term and decreased after prolonged, i.e., lasting several hours, exercise (Viru and Viru, 2004). It is, however, difficult to distinguish the exercise-related (physiological) from the emotional related (psychological) stress induced by extrinsic factors that are dependent on intrinsic factors, such as age or experience and temperament, which might be better represented by ocular temperature (Negro *et al.*, 2018).

9.2 Eye Temperature Measured by Infra-Red Thermography

The rise of the eye caruncula temperature measured by infrared thermography was reported as a reliable indicator of stress in animals. It was often studied together with salivary cortisol measurements to assess the stress response to transportation (Schmidt *et al.*, 2010a) and husbandry procedures (Yarnell *et al.*, 2013). Equestrian practices deemed to cause discomfort to the horse, such as neck hyperflexion (Hall *et al.*, 2014) or a tight noseband (Fenner *et al.*, 2016), were also studied. It is believed that the rise in eye temperature represents an emotional response to stressors, including exercise, i.e. a measure of emotional reactivity to effort, that can have a beneficial or detrimental effect on performance. In the same line of this research, eye temperature was recently proposed as a selection tool to help identify emotional reactivity as a desirable, or undesirable, a trait to performance according to the intended use of the horse (Negro *et al.*, 2018; Sánchez *et al.*, 2016).

Like SC, the rise of eye temperature is believed to result from the activation of the HPA axis, which ensures an increase of the periorbital blood flow in stressed animals (König v. Borstel *et al.*, 2017). One of the proposed added values of the use of eye temperature is its potential independence of the effort effect and, thus, a valid mean of evaluating emotional stress in exercised horses (König v. Borstel *et al.*, 2017). In sport horses, eye temperature was investigated in showjumping (Bartolome et al., 2013; Valera et al., 2012) and dressage competitions (Sánchez, Bartolomé, & Valera, 2016), in Standardbred harness races (Negro et al., 2018), in flat race Arabian and Thoroughbred horses in training (Soroko et al., 2016), but nor endurance.

9.3 Future Perspectives

It is still challenging to untangle emotional distress and experienced pain from the effort stressor in the exercising horse. As the scientific community has recognised these limitations, there has been a shift in the last years investigating behavioural indicators of distress due to pain, such as the grimace score and conflict behaviours. An attractive, innovative approach is artificial intelligence through video analysis of facial pain expression to assess animal welfare through physical manifestations (Andersen, 2018).

Compared with horse racing, endurance riding research to prevent severe and catastrophic injuries is still in infancy. Despite data-driven rule changes, as it is currently taking place in endurance, other efficient and affordable non-invasive biomarkers are being investigated (Page *et al.*, 2021). Omics, or the study of protein, genetic material (both DNA and RNA, including microRNAs—small non-coding ribonucleic acids) and metabolites, is a new research area whose value in fracture prediction has been investigated in the last years in horses and humans (Lee *et al.*, 2021). Tested on blood, and thus invasive, the expression of mRNA studies in horseracing (Page *et al.*, 2021) and in endurance (Mach *et al.*, 2016) have not yet proven its efficiency.

10. Objectives

Non-invasive biomarkers indicative of pain, if used concomitantly with inertia sensor-based devices to quantify locomotion, could help target more efficiently non-fit horses to compete in endurance and reduce unnecessary eliminations.

The general goal of this thesis was to explore the potential usefulness of non-invasive techniques in the gait and welfare assessment of endurance horses. The specific aims of each of the performed studies were:

1.) To document the necessity of this study by characterising the difficulties encountered by FEI Endurance Official Veterinarians (EOV) in gait evaluation during endurance competitions, the frequency of rider confrontations after deciding to eliminate a horse and the receptiveness of EOV's to new objective technologies to quantify lameness.

2.) To start assessing the feasibility of detecting and quantifying gait during endurance competitions with an inertial sensor-based system.

3.) To investigate the impact of a time delay necessary to instrument a horse with sensors for gait analysis on the Cardiac Recovery Index (CRI).

4.) To determine the impact of two different trotting presentation styles, e.g., a regular trot-up as for a lameness workup and a trot-up as interpreted by handlers in a competition scenario at a vet gate, in the subjective scores and agreement of FEI endurance veterinarians

5.) To determine trends in salivary cortisol and eye temperature measured by infrared thermography and its variation before and during endurance

References:

AERC. (2008, 11/08). Endurance Rider’s Handbook. 3.0 Rev. Retrieved from *https://aerc.org/static/AERC_Rider_Handbook.pdf*

Aubets, J., & Segura, J. (1995). Salivary cortisol as a marker of competition related stress. *Science & Sports, 10*(3), 149-154. doi:*https://doi.org/10.1016/0765-1597(96)89361-0*

Bartolomé, E., & Cockram, M. S. (2016). Potential Effects of Stress on the Performance of Sport Horses. *Journal of Equine Veterinary Science, 40*, 84-93. doi:*https://doi.org/10.1016/j.jevs.2016.01.016*

Bartolome, E., Sanchez, M. J., Molina, A., Schaefer, A. L., Cervantes, I., & Valera, M. (2013). Using eye temperature and heart rate for stress assessment in young horses competing in jumping competitions and its possible influence on sport performance. *Animal, 7*(12), 2044-2053. doi: *https://doi.org/10.1017/S1751731113001626*

Becker-Birck, M., Schmidt, A., Lasarzik, J., Aurich, J., Möstl, E., & Aurich, C. (2013). Cortisol release and heart rate variability in sport horses participating in equestrian competitions. *Journal of Veterinary Behavior: Clinical Applications and Research, 8*(2), 87-94. doi:*https://doi.org/10.1016/j.jveb.2012.05.002*

Bennet, E. D., Parkin, T. D. H. (2017). *Risk Factors in FEI Endurance Rides 2010-2016*. Paper presented at the Endurance Forum, Vic, Barcelona, Spain. *https://inside.fei.org/content/endurance-fei-endurance-forum-presentations-1*

Bohák, Z., Szenci, O., Harnos, A., Kutasi, O., & Kovács, L. (2017). Effect of temperament on cortisol response to a single exercise bout in Thoroughbred racehorses - short communication. *Acta Vet Hung, 65*(4), 541-545. doi:*https://doi.org/10.1556/004.2017.052*

Cayado, P., Muñoz-Escassi, B., Domínguez, C., Manley, W., Olabarri, B., Sánchez de la Muela, M., . . . Vara, E. (2006). Hormone response to training and competition in athletic horses. *Equine Vet J Suppl*(36), 274-278. doi:*https://doi.org/10.1111/j.2042-3306.2006.tb05552.x*

Contreras-Aguilar, M. D., Martínez-Subiela, S., Cerón, J. J., Martín-Cuervo, M., Tecles, F., & Escribano, D. (2019). Salivary alpha-amylase activity and concentration in horses with acute abdominal disease: Association with outcome. *Equine Vet J, 51*(5), 569-574. doi:*https://doi.org/10.1111/evj.13066*

Cook, N., Schaefer, A., Warren, L., Burwash, L., Anderson, M., & Baron, V. (2001). Adrenocortical and metabolic responses to ACTH injection in horses: an assessment by salivary cortisol and infrared thermography of the eye. *Can J Anim Sc, 81*.

Davidson, E. J., & Ross, M. W. (2003). Clinical recognition of stress-related bone injury in racehorses. *Clinical Techniques in Equine Practice, 2*(4), 296-311. doi:*https://doi.org/10.1053/j.ctep.2004.04.002*

de Graaf-Roelfsema, E., Keizer, H. A., van Breda, E., Wijnberg, I. D., & van der Kolk, J. H. (2007). Hormonal responses to acute exercise, training and overtraining a review with emphasis on the horse. *Veterinary Quarterly, 29*(3), 82-101. doi:*https://doi.org/10.1080/01652176.2007.9695232*

de Mira M.C., Santos, C., Lopes, M. A., & Marlin, D. J. (2019). Challenges encountered by Federation Equestre Internationale (FEI) veterinarians in gait evaluation during FEI endurance competitions: an international survey. *Comparative Exercise Physiology, 15*(5), 371-378. doi:*https://doi.org/10.3920/cep180058*

de Mira, M. C., Williams, J., Santos, R. G. d., Rodrigues, P., Arroja, B., & Marlin, D. (2020). Do waiting times in endurance vet gates affect the cardiac recovery index? *Comparative Exercise Physiology*, 1-8. doi:*https://doi.org/10.3920/CEP190081*

Desmecht, D., Linden, A., Amory, H., Art, T., & Lekeux, P. (1996). Relationship of plasma lactate production to cortisol release following completion of different types of sporting events in horses. *Vet Res Commun, 20*(4), 371-379.

FEI. (2020). FEI History. Retrieved from *https://inside.fei.org/fei/about-fei/history*

FEI. (2020a, 1st July 2020). Endurance Rules. *11th edition. Updates effective 1 Jan 2020.* 11th. Retrieved from *https://inside.fei.org/sites/default/files/FEI%20Endurance%20Rules%20-%201%20July%202020%20-%2016.12.2019%20-%20Clean.pdf*

FEI. (2020b, 1 Jan 2020). Veterinary Regulations. 14th Edition. Retrieved from *https://inside.fei.org/fei/regulations/veterinary*

Fenner, K., Yoon, S., White, P., Starling, M., & McGreevy, P. (2016). The Effect of Noseband Tightening on Horses' Behavior, Eye Temperature, and Cardiac Responses. *PLoS One, 11*(5), e0154179. doi:*https://doi:10.1371/journal.pone.0154179*

Hall, C., Kay, R., & Yarnell, K. (2014). Assessing ridden horse behavior: Professional judgment and physiological measures. *Journal of Veterinary Behavior: Clinical Applications and Research, 9*(1), 22-29. doi:*https://doi.org/10.1016/j.jveb.2013.09.005*

Hardeman, A., Serra Bragança F., Swagemakers J.H., Van Weeren, R., Roepstorff, L., Koene M. (2018). *Variation in gait symmetry parameters in sound horses at trot on the straight line and on the lunge*. Paper presented at the International Conference on Equine Exercise Physiology (ICEEP), Lorne, Australia. *https://doi.org/10.3920/cep2018.s1*

Hyyppä, S. (2005). Endocrinal responses in exercising horses. *Livestock Production Science, 92*(2), 113-121. doi:*https://doi.org/10.1016/j.livprodsci.2004.11.014*

Janczarek, I., Bereznowski, A., & Strzelec, K. (2013). The influence of selected factors and sport results of endurance horses on their saliva cortisol concentration. *Pol J Vet Sci, 16*(3), 533-541.doi: *https://10.2478/pjvs-2013-0074*

Kędzierski, W., & Cywińska, A. (2014). The Effect of Different Physical Exercise on Plasma Leptin, Cortisol, and Some Energetic Parameters Concentrations in Purebred Arabian Horses. *Journal of Equine Veterinary Science, 34*(9), 1059-1063. doi:*https://doi.org/10.1016/j.jevs.2014.06.005*

Keegan, K. G. (2011). Objective assessment of lameness. In G. M. Baxter (Ed.), *Adams & Stashak's Lameness in Horses* (6 ed., pp. 154-164). Oxford: Willey-Blackwell.

König v. Borstel, U., Visser, E. K., & Hall, C. (2017). Indicators of stress in equitation. *Applied Animal Behaviour Science, 190*, 43-56. doi: *https://doi.org/10.1016/j.applanim.2017.02.018*

Lee, E. C., Fragala, M. S., Kavouras, S. A., Queen, R. M., Pryor, J. L., & Casa, D. J. (2017). Biomarkers in Sports and Exercise: Tracking Health, Performance, and Recovery in Athletes. *J Strength Cond Res, 31*(10), 2920-2937. doi:*https://doi.org/10.1519/JSC.0000000000002122*

Lopes, M. A. F., Eleuterio, A., & Mira, M. C. (2018). Objective Detection and Quantification of Irregular Gait With a Portable Inertial Sensor-Based System in Horses During an Endurance Race—a

Preliminary Assessment. *Journal of Equine Veterinary Science, 70*, 123-129. doi:*https://doi.org/10.1016/j.jevs.2018.08.008*

Marlin, D., & Williams, J. (2018). Equine endurance race pacing strategy differs between finishers and non-finishers in 120 km single-day races. *Comparative Exercise Physiology, 14*(1), 11-18. doi:*https://doi.org/10.3920/cep170027*

Milgrom, C., Giladi, M., Stein, M., Kashtan, H., Margulies, J. Y., Chisin, R., . . . Aharonson, Z. (1985). Stress fractures in military recruits. A prospective study showing an unusually high incidence. *J Bone Joint Surg Br, 67*(5), 732-735. doi:*https://doi.org/10.1302/0301-620X.67B5.4055871*

Misheff, M. M., Alexander, G. R., & Hirst, G. R. (2010). Management of fractures in endurance horses. *Equine Veterinary Education, 22*(12), 623–630. doi:*https://10.1111/j.2042-3292.2010.00150.x*

Nagy, A., Murray, J. K., & Dyson, S. J. (2014). Descriptive epidemiology and risk factors for eliminations from Federation Equestre Internationale endurance rides due to lameness and metabolic reasons (2008-2011). *Equine Vet J, 46*(1), 38-44. doi:*https://doi.org/10.1111/evj.12069*

Negro, S., Bartolomé, E., Molina, A., Solé, M., Gómez, M. D., & Valera, M. (2018). Stress level effects on sport performance during trotting races in Spanish Trotter Horses. *Res Vet Sci, 118*, 86-90. Doi:*https://doi.org/10.1302/0301-620X.67B5.405587110.1016/j.rvsc.2018.01.017*

Nunes, J. A., Crewther, B. T., Ugrinowitsch, C., Tricoli, V., Viveiros, L., de Rose, D., Jr., & Aoki, M. S. (2011). Salivary hormone and immune responses to three resistance exercise schemes in elite female athletes. *J Strength Cond Res, 25*(8), 2322-2327. doi: *https://doi.org/10.1519/JSC.0b013e3181ecd033*

Peeters, M., Closson, C., Beckers, J.-F., & Vandenheede, M. (2013). Rider and Horse Salivary Cortisol Levels During Competition and Impact on Performance. *Journal of Equine Veterinary Science, 33*(3), 155-160. doi: *https://doi.org/10.1016/j.jevs.2012.05.073*

Peeters, M., Sulon, J., Beckers, J. F., Ledoux, D., & Vandenheede, M. (2011). Comparison between blood serum and salivary cortisol concentrations in horses using an adrenocorticotropic hormone challenge. *Equine Vet J, 43*(4), 487-493. doi: *https://10.1111/j.2042-3306.2010.00294.x*

Redaelli, V., Luzi, F., Mazzola, S., Bariffi, G. D., Zappaterra, M., Nanni Costa, L., & Padalino, B. (2019). The Use of Infrared Thermography (IRT) as Stress Indicator in Horses Trained for Endurance: A Pilot Study. *Animals (Basel), 9*(3). doi: *https://doi.org/10.3390/ani9030084*

Riber, C., Cuesta, I., Munoz, A., Gata, J., Trigo, P., & Castejon, F. M. (2006). Equine locomotor analysis on vet-gates in endurance events. *Equine Vet J Suppl*(36), 55-59. doi: *https://doi.org/10.1111/j.2042-3306.2006.tb05513.x*

Sánchez, M. J., Bartolomé, E., & Valera, M. (2016). Genetic study of stress assessed with infrared thermography during dressage competitions in the Pura Raza Español horse. *Applied Animal Behaviour Science, 174*, 58-65. doi: *https://doi.org/10.1016/j.applanim.2015.11.006*

Sloet Van Oldruitenborgh-Oosterbaan, M. M. (2018). *Subjective and objective evaluations of horses for a fit-to-compete or unfit-to-compete judgement*. Paper presented at the 10th International Conference on Equine Exercise Physiology, Lorne, Australia.

Soroko, M., Howell, K., Zwyrzykowska, A., Dudek, K., Zielińska, P., & Kupczyński, R. (2016). Maximum Eye Temperature in the Assessment of Training in Racehorses: Correlations With Salivary Cortisol Concentration, Rectal Temperature, and Heart Rate. *Journal of Equine Veterinary Science, 45*, 39-45. doi: *https://doi.org/10.1016/j.jevs.2016.06.005*

Stewart, M., Wilson, M. T., Schaefer, A. L., Huddart, F., & Sutherland, M. A. (2017). The use of infrared thermography and accelerometers for remote monitoring of dairy cow health and welfare. *Journal of Dairy Science, 100*(5), 3893-3901. doi:*https://doi.org/10.3168/jds.2016-12055*

Symons, F. J., Eighazi, I., Reilly, B. G., Barney, C. C., Hanson, L., Panoskaltsis-Mortari, A., Wilcox, G. L. (2015). Can Biomarkers Differentiate Pain and No Pain Subgroups of Nonverbal Children with Cerebral Palsy? A Preliminary Investigation Based on Noninvasive Saliva Sampling. *Pain Medicine, 16*(2), 249-256. doi:*https://doi.org10.1111/pme.12545*

Valera, M., Bartolomé, E., Sánchez, M. J., Molina, A., Cook, N., & Schaefer, A. (2012). Changes in Eye Temperature and Stress Assessment in Horses During Show Jumping Competitions. *Journal of Equine Veterinary Science, 32*(12), 827-830. doi:*https://doi.org/10.1016/j.jevs.2012.03.005*

Viry, S., 2014. Etude préliminaire sur l'influence de la monte « styledésert» sur le couplage cavalier-cheval en course d'endurance. In: I.f.d.c.e.d. l'équitation (Ed.), 40ème Journée de la Recherche Équine, March, 18th Paris. I.F.C.E.

CHAPTER II: Challenges Encountered by Federation Equestre Internationale (FEI) Veterinarians in Gait Evaluations during FEI Endurance Competitions: An International Survey

Mónica C. de Mira, Carina Santos, Marco A.F. Lopes, David Marlin (2019). Challenges Encountered by Federation Equestre Internationale (FEI) Veterinarians in Gait Evaluations during (FEI) Endurance Competitions: An International Survey. Comparative Exercise Physiology. 2019;15(5):371-8. https://doi.org/10.3920/CEP180058

Comparative Exercise Physiology, 2019; 15 (5): 371-378

Wageningen Academic Publishers

ICEEP 2018

Challenges encountered by Federation Equestre Internationale (FEI) veterinarians in gait evaluation during FEI endurance competitions: an international survey

M.C. de Mira[1*], C. Santos[2], M.A. Lopes[3] and D.J. Marlin[4]

[1]Instituto de Ciências Agrárias e Ambientais Mediterrânicas-ICAAM, Universidade de Évora, Núcleo da Mitra Apartado 94, 7006-554 Évora, Portugal; [2]Iberovet, Rua do Bocage, Lote 32, R/C Dto. Portugal, 2890-052 Alcochete, Portugal; [3]School of Animal and Veterinary Sciences, University of Adelaide, Equine Health & Performance Centre, 1454 Mudla Wirra Rd, 5371 Roseworthy, SA, Australia; [4]David Marlin Consulting, P.O. Box 187, Cambridge, United Kingdom; monicademira@gmail.com

Received: 7 December 2018 / Accepted: 12 June 2019

RESEARCH ARTICLE

Abstract

Equine endurance competitions are long races over 1-2 days and horses can be eliminated in international competitions by FEI veterinarians for lameness at any time. Elimination rates due to lameness are high, affect the outcome of the races and commonly elicit objections to the decision from endurance competitors. The aims of this study were: (1) to assess the opinion of FEI official endurance veterinarians (OEVs) about the challenges in assessing lameness; (2) the occurrence of confrontations with riders when horses are eliminated from races due to lameness; (3) to assess OEVs' thoughts about the adoption of user-friendly technology for objective gait evaluation to help detect, quantify and document lameness. All FEI OEVs were asked to complete a questionnaire. There were 157 responses, being most of the respondents from Europe (56.1%) or the Middle East (16.6%). For the majority of respondents, detection of lameness was considered challenging, even for experienced and well-trained veterinarians (57.3%). OEVs also considered it was often hard to classify horses as lame or sound (65.8%). Handlers not trotting the horse appropriately during gait evaluations was considered the most common problem compromising the evaluation (94.3%). Most OEVs (98.2%) responded that they had been confronted at least once by a rider or associate about the decision to eliminate a horse due to lameness. Most OEVs (71.3%) would be interested in having the support of user-friendly technology for objective gait evaluation (33.3% for all evaluations; 38.0% only when horses have more subtle gait abnormalities). The findings of this survey suggest that technology to objectively detect and quantify gait abnormalities during endurance competitions would be beneficial to support decisions made by OEVs when evaluating and or eliminating horses for lameness.

Keywords: lameness, objective, sport, quantification, sensor

1. Introduction

Gait assessment is a major component of veterinary evaluations performed during endurance competitions and is considered paramount for the well-being of the horses. Endurance competitions are long races over 1 or 2 days and horses can be eliminated by Fédération Equestre Internationale (FEI) Official Endurance Veterinarians (OEVs) at any time, even after crossing the finish line (FEI, 2018a,c). OEVs are accredited by the FEI to work in endurance competitions and are rated by a star-system level (2-star lowest level, 3-star intermediate level, or 4-star highest level) according to their experience and mandatory training (FEI, 2018a). Due to logistical and time constraints, lameness during endurance competitions is currently limited to inspection of the horse trotted back and forth in hand on a 30-50 m straight line (FEI, 2018c; Holbrook, 2011; Misheff, 2011). According to the FEI veterinary rules, OEVs are expected to determine which irregular gaits cause pain or threaten the immediate ability of the horse to safely perform athletically (FEI, 2018a). Even if it is common sense that to subject a lame horse to intense and or prolonged exercise may aggravate an existing lesion(s) or lead to additional lesions (Baxter and

ISSN 1755-2540 print, ISSN 1755-2559 online, DOI 10.3920/CEP180058

Stashak, 2011; Holbrook, 2011; Misheff, 2011; Ross, 2011), a mild lameness might nevertheless be acceptable. Contrary to human athletes, which can self-determine their limits and what their discomfort limit is, for veterinarians to establish whether horses are experiencing pain, or not, is a task far from trivial (Van Weeren *et al.*, 2017). The elimination rates for lameness account to near, or as high as, 70% of all eliminations (Coombs and Fisher, 2012; Marlin and Williams, 2018; Nagy *et al.*, 2010, 2012, 2013, 2014a,b; 2017; Younes *et al.*, 2016). On the competitors side a failure to qualify due to lameness affects the outcome of the competitions, as well as the market value of a horse, may lead to failure to secure a prize money, prolongs the qualification process of a horse and commonly elicits complaints from endurance competitors for its subjectivity (Nagy *et al.*, 2012). Until recently the use of objective gait analysis technology for lameness detection outside the laboratory was impractical and cost-prohibitive, but this paradigm has been changed by new technologies that are relatively quick, technically undemanding and inexpensive (Keegan, 2011; Riber *et al.*, 2006). The aim of the present study was to characterise the difficulties encountered by FEI OEVs in gait evaluation during endurance competitions, the frequency of rider confrontations after a decision to eliminate a horse and the receptiveness of OEV's to new objective technologies to quantify lameness.

2. Material and methods

This study was a cross-sectional survey that used a convenience sample. The survey comprised a questionnaire to self-complete in English circulated to 221 FEI official endurance veterinarians (OEVs) between 2013 and 2017, either by hand by the first author during international competitions (n=100) or sent by email (n=121) as a Google docs form (see Supplementary material). The majority of the respondents were from Europe (46.5%, n=73) followed by the Gulf Countries (16.6%, n=26), with 39 different nationalities represented. Although not anonymous to the first author, the questionnaire was confidential. The questionnaire was divided in three main parts. The first part concerned the level and experience of the OEVs and comprised questions requiring short answers and two multiple choice questions. The second part specifically regarded the various difficulties that can be encountered by OEVs during lameness evaluations, including confrontation by the rider, as a dichotomic multiple choice (yes/no). Optic illusion created by a pelvic asymmetry, i.e. asymmetric positioning of one of the tuber coxae caused by a fracture and/or correspondent muscle atrophy, confounding hindlimb lameness, was also included. The frequency of occurrence was categorised as 4 multiple choice options (very common, common, very uncommon, never seen). The third part aimed to evaluate the receptiveness of OEVs to new technologies to objectively assess lameness using one multiple-choice question and one open-ended (free-text) question.

The data were analysed in SPSS (IBM SPSS Statistics for Windows, Version 22.0; IBM Corp., Armonk, NY, USA). Descriptive statistics were used to characterise the sample of OEVs according to their experience as equine veterinarians, in endurance and as FEI veterinarians, FEI star level, country registered and respective geographical location. The OEV lists of the starting (2013) and end (2017) points of the study were used to compare the sample with the population. To assess significant differences (star level and National Federation distribution) a Chi Square Test was used. Descriptive statistics were used to characterise the multiple-choice answers and in order to infer associations with experience and star-level of the OEVs, proportion or a Chi Square, or, when assumptions were not met, Kruskal-Wallis or a Fisher Exact tests were used. Significance was set at $P<0.05$.

3. Results

Of the 221 OEVs contacted, the overall response rate was 71% (n=157). The response rate for surveys given by hand was 97% (n=97) and 50% (n=60) for the survey sent by email. Regarding the star level of the respondents, 24.8% (n=39) of the OEVs were classified as 2-star, 33.8% (n=53) as 3-star and 41.4% (n=65) as 4-star. The participants mean number of years of experience respectively as equine veterinarians, endurance veterinarians and as years as FEI veterinarians was 20.3±10.5(SD) years, 16.4±9 and 11±8 years, respectively. As expected, there was a positive association ($P<0.001$) between the mean number of years of experience, years of experience as endurance veterinarian and FEI veterinarian with star-level classification. When OEVs were asked on their opinion about detecting lameness during an endurance competition, over a half (57.3%, n=90) considered that, even for an experienced and well-trained veterinarian, it is a challenging work. Close to a third (31.2%, n=49) thought it was a straightforward task and only a minority (10.2%, n=16) claimed that lameness detection can be performed by any veterinarian after a minimal training. A relationship between star-level and the different answers was not found to be significant ($P=0.678$). When asked about their own ability in detecting lameness during endurance competitions, two thirds of the respondents (65.8%, n=102) thought that, although they are experienced veterinarians, many times it was hard for them to define if a horse was lame or not during endurance races, over a quarter (28.4%, n=44) answered that they did not have any trouble detecting lameness during endurance competitions with a larger proportion of 4-star OEVs (31,3%). Only 6% (n=9) expressed that they were 'still learning.' There were not enough counts to establish a relationship with star level for this answer. The factors considered by OEVs to compromise their ability to detect lameness and their frequency of occurrence are summarised in Figure 1. An inappropriate presentation by hand by the rider or associate was the most important factor considered by OEVs to affect gait assessment with 94% (n=148) responding it compromises their ability to detect lameness during endurance competitions. This was as a common occurrence for two thirds (67% n=99) of the respondents. Bad horse behaviour and an inappropriate alley for lameness evaluation ranked second and third, but were considered by the majority to be uncommon. Poor lighting and a heavy rainfall also were considered to affect lameness detection, but only occurring uncommonly. Even if uncommon for most, optical illusions created by pelvic asymmetry was considered by more than half to be a disturbing factor for lameness detection. Regarding an association between star-level and experience, there was no significant difference for all factors except pelvic asymmetry and poor lighting. OEVs with more years as endurance (P=0.042) or as FEI veterinarians (P=0.039) were less likely to consider optical illusions created by pelvic asymmetry a relevant factor. On the other hand, OEVs with more experience as equine practitioners (P<0.01) and as endurance veterinarians (P=0.004) considered unsatisfactory lighting as an important factor that influenced the ability of lameness detection. Concerning altercations with riders or associates over a lameness related elimination, the vast majority of the OEVs (94.9%, n=148) claimed to have had experienced it with 36% (n=56) reporting this to be common. The star level or experience did not significantly affect the response. When questioned if they would be receptive to a technology, that did not interfere with their pace of work, to help objectively detect and quantify lameness and document the results of the evaluations during an endurance competition, over two thirds of the respondents (71.3% n=107) agreed, one third did not see any benefit (18%, n=27) and the remaining chose 'other answer' (10.7%, n=16).

4. Discussion

This survey intended to characterise the challenges faced by veterinarians performing lameness evaluations during endurance competitions and receptiveness to new technologies to evaluate gait. Until now, most reports about factors affecting lameness assessment have been typically limited to personal observations, described in lameness textbooks (Baxter and Stashak, 2011; Ross, 2011) or review papers (Dyson, 2014). To our knowledge, identification of those factors which are more likely to affect lameness detection from the perspective of a large group of veterinarians has not been studied before. The results of this questionnaire, and ultimately the introduction of new technologies, might not only help generate better guidelines to ameliorate the non-quantifiable nature of subjective gate assessment, but also facilitate early identification of severe injuries, enhance accuracy and consistency of OEVs judgements, and, therefore, improve the perception from the competitor's perspective that the decision to eliminate a horse is fair.

From a list of factors, based on from previous reports and first author's experience, OEVs across all star-levels participating in this survey identified an inappropriate presentation by the handler as the most important and common factor affecting lameness detection during endurance competitions. Due to its importance to the outcome of a competition, and implicitly to the success of the triad rider/trainer/owner, trot-ups are extensively trained before the competitions to minimise any suspicion of lameness. Some commonly observed techniques can be elevating and holding the head with a short lead rope grip (mainly when trotting away from the veterinarian and particularly during deceleration), an abrupt deceleration to avoid lame steps, bending the neck, pulling the horse, waving the lead rope, increasing or decreasing the speed of

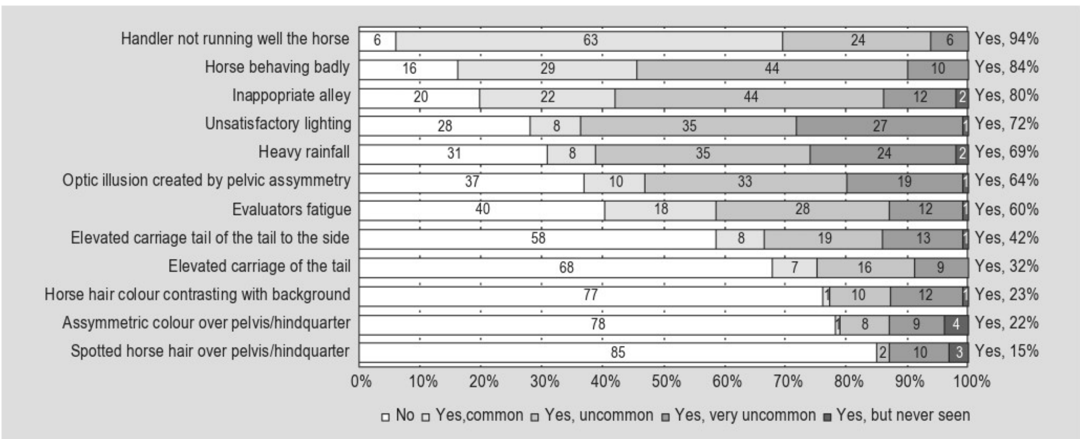


Figure 1. Results of the factors that might compromise the ability to detect lameness and frequency of occurrence.

trot to its minimum and obstructing the view of the OEV by running in the front of the horse (M. Mira, personal communications). Speed of trot might be changed, intentionally or unintentionally, by the handler and may influence the degree of lameness(Starke *et al.*, 2013). A constant change in trot speed can make gait evaluation difficult, but the deceleration can reveal useful information in subtle lameness (Baxter and Stashak, 2011; Ross, 2011). Mildly lame horses undergoing subjective clinical examination were shown to be more likely declared sound when trotted at higher speeds (Starke *et al.*, 2013). Lameness textbooks suggest to trot horses as slowly as practical (Baxter and Stashak, 2011), or alternatively, not too slow and not too fast (Ross, 2011). Anyhow, a control of speed has been recommended to maintain consistency of results in both subjective and objective gait assessment in repeated measurements (Serra Braganca, 2018; Starke *et al.*, 2013). The FEI rules are presently limited to instruct handlers to lead the horse from the left hand side on a loose rein and in the centre of the inspection track in a flat firm surface(FEI, 2018c). More specific written guidelines including speed of trot and indicating procedures that are not allowed could ameliorate the uniformity and consistency of trot-ups.

Respondents to this survey ranked behaviour of the horse and an inappropriate trot-lane as the next most important factors to affect lameness detection. However, the majority thought it occurred uncommonly, most likely because of the usual good-tempered nature of endurance horses, with the exception of some young horses in novice competitions, and the existence of permanent venues for higher level endurance competitions or the modernisation of many competition sites making an inappropriate lane a rarer occurrence. An interesting finding was the significant trend for veterinarians with more years as equine and endurance veterinarians to consider poor lighting a relevant factor to affect lameness detection. This might be explained by older veterinarians having reduced visual acuity, or being more confident to admit poor lighting as a limitation. There was a trend for OEVs with less years of experience as endurance or as FEI veterinarians to think that optical illusions created by a pelvic asymmetry were a confounding factor for hindlimb lameness evaluation. Unilateral gluteal muscle atrophy or asymmetry of the *tubera sacrale* had been previously described as confounding factors for hindlimb gait assessment (Dalin *et al.*, 1985; Dyson, 2009). Also, hindlimb assessment in mild lameness was considered to be less reliable (Dyson, 2009; Parkes *et al.*, 2009; Starke and Oosterlinck, 2019). In fact, high level competitors tend to avoid selecting noticeably asymmetric horses as they perceive occurrence of subjectivity by the OEVs in the gait evaluation of such horses, and if they choose to present such a horse, they will often point out the asymmetry to the veterinarian in the lane beforehand (M. Mira, personal communications). Although considered an uncommon occurrence, almost half of the respondents identified an uneven carriage of the tail blocking the view of one side of the pelvis as a potential limiting factor such as reported before (Dyson, 2009; Ross, 2011).

Evaluator fatigue at the end of the day and/or after examining a large number of horses was considered to be a relevant factor for the large majority of the respondents of this survey. Endurance competitions can start very early in the morning, or even during the night, precluding a good night sleep for most of the OEVs at competitions. Taking in consideration the length of most international competitions and that long hours and repetitive tasks combined are known to induce fatigue (Tomei *et al.*, 2006), it was surprising that most considered fatigue an uncommon occurrence. This might show a reluctance or prejudice of OEVs to assume a weakness regarding their resilience to adversity.

The limited tools that OEVs dispose of, regarding the complexity of a lameness exam and the severity of injuries that might occur at the present times in competitions, the scarce and vague published information defining a 'non-fit-to compete' lameness by the FEI and the reported subjectivity of mild lameness examination among veterinarians makes lameness detection during endurance competitions a challenging task, regardless of star-level. Therefore, it's not surprising that the majority of the respondents, even if they considered themselves very experienced, still thought that for them it was difficult many times to define if a mild lameness was significant or not. Mild lameness identification has been recognised by multiple authors to be a difficult task, showing lack of sensitivity and repeatability even when veterinarians are not working under pressure and time constraints such as in an endurance competitions (Fuller *et al.*, 2006; Hewetson *et al.*, 2006; Ishihara *et al.*, 2005; Keegan *et al.*, 1998, 2010; McCracken *et al.*, 2012). Supporting this was a recent study involving experienced FEI veterinarians that showed, not only a poor inter-, but also a meagre intra-agreement of 'fit-to-compete' versus 'non-fit-to compete' when judging a mild lameness of the same horses three months later by video-analysis (Sloet Van Oldruitenborgh-Oosterbaan, 2018). Furthermore, most ancillary tests thought to be essential for a complete lameness examination are impractical for use during endurance competitions due to time, cost and venue setting (FEI, 2018c; Nagy *et al.*, 2014a). Therefore, common findings such as a bilateral fore or hindlimb, or a coexistent forelimb plus hindlimb lameness, will not be clearly identified in a straight line (Greve and Dyson, in press). For instance, stress fractures might occur bilaterally and one of the most consistent findings is resentment to palpation of the fractured region (MacKinnon *et al.*, 2015; Pleasant *et al.*, 1992). This procedure is currently discouraged during veterinary inspections (FEI, 2018c), even if the FEI recently introduced an update on flagged horses (that suffered four consecutive eliminations for lameness in national and international events) that instruct the holding lane veterinarian to evaluate those horses in the circle and to palpate relevant tissues (FEI, 2018b). The simplistic approach of the FEI veterinary rules to solely define an unsound 'non-fit-to-compete' as an irregular gait, which must be consistently observable at trot, leaves room to a 'grey zone' of interpretations by OEVs and competitors. Horses can be presented with a consistent but very mild lameness, which might be of limited importance in relation to some horses with some clearly lame, but not, consistent steps in a trot-up, many times further complicated by handler manoeuvres (M. Mira, personal communication). Furthermore, to our knowledge, there is no written support to the A,B and C grading system used in competitions, except for a document sent to OEVs in 2013 when the Global Injuries Surveillance was implemented (FEI, 2013). The American Endurance Ride Conference (AERC) rules, that represent the foundation of equestrian endurance rules worldwide, are far more detailed regarding lameness assessment. First the AERC uses the lameness scale of the Association of American Equine Practioners, which is the most familiar system to equine practitioners worldwide, to define a 'non fit-to-compete' lameness. Secondly, even if a grade III matches the FEI definition of unfit to compete, some important caveats are made. Not only it is clearly stated that a horse that only takes one or two questionable steps can be eliminated if the control judge feels that continuing the ride could cause irreparable damage to the horse, but also the reserve is made that a horse that shows a consistent slight limp might be allowed to continue if the lameness is clearly due to a superficial injury, loss of a shoe, or some other temporary and relatively insignificant factor. Therefore, in doubtful lameness it is encouraged to circle the horse, to perform detailed palpation, flexion tests, hoof testers and other diagnostic procedures to help to segregate the innocently sore from the dangerously lame (AERC, 2008a,b; Mackay-Smith, 2016). An ongoing debate about the dichotomy of sound *versus* lame and unfitness of lame horses to compete has been taking place in the editorial section of a journal with a large media coverage (Adair *et al.*, 2018, 2019; Bathe *et al.*, 2018; Dyson, 2019; Van Weeren *et al.*, 2017, 2018). The discussion was triggered by the reports of a large percentage of performant and owner-sound horses with no lameness complaints found to be lame, both, in a subjective in-depth lameness work-up (Dyson and Greve, 2016) and in an objective assessment using a sensor-based system objective evaluation (Rhodin *et al.*, 2017). Although it is difficult to establish what is the threshold for an acceptable gait asymmetry it might also be unrealistic to expect that a high competition athlete will always be sound during his sporting career. Even if horses might reasonably be expected to perform better if free from lameness (Dyson and Greve, 2016; Parente *et al.*, 2002), still some lame horses may also perform better than sound ones (Keegan, 2007).

The high number of OEVs, regardless of star-level or experience, having experienced a confrontation in this survey reflect, not only, the high frequency of eliminations for lameness, but also the non-measurable subjectivity of gait assessment perceived by riders and associates. It had been previously reported that eliminations for lameness commonly elicit complaints from endurance competitors (AERC, 2008a; Nagy *et al.*, 2012). Even though the FEI introduced the rule that a panel of three OEVs has to be consulted, each veterinarian voting anonymously, before eliminating any horse (FEI, 2018c) to remove the onus of an elimination of a particular veterinarian and to make decisions more consistent, it still falls on the individual OEV to decide whether a second evaluation by the panel is needed or not and therefore this did not completely solve the problem. Ideally a panel would evaluate each trot-up, but this is not feasible with the current set up of a vet gate.

Considering the points discussed above and that whether to eliminate a horse or not often poses a moral distress for the OEVs, it is not surprising that the majority of the respondents to this survey were receptive to a new technology that would objectively measure lameness and help them take better decisions. The feasibility of the use of objective gait technology using a portable inertial sensor-based system in horses during an endurance race has been assessed previously (Lopes *et al.*, 2018). However, although it was previously reported that veterinarians and riders or associates were not resistant to the time taken to instrument horses, the study was performed during a controlled speed qualifying competition and most likely might not be applicable to a more competitive scenario. Time to instrument horses was not registered in this study, but it has been previously reported to take an average of less than three minutes per horse (Keegan *et al.*, 2011). Even if this extra-time could be reduced with trained personnel, it would currently be hard to fit in an already busy vet gate, not to mention the required additional time to interpret results in real time. The importance of lameness detection during endurance competitions, with regard to horse welfare and race outcome, and the limitations of gait assessment based exclusively on a brief subjective examination would suggest that the use of an objective gait analysis system would likely benefit all segments of the equine endurance industry, yet there are still some important limitations. Several years of research and many starters in endurance competitions would be needed to determine a cut-off value for an acceptable 'fit-to-compete' gait in an endurance competition context. At the most, and in a first approach, one could envision that such a system would be solely used to objectively compare the gait of a horse to himself throughout a competition. Yet, further investigations are needed to determine if there is sufficient repeatability of a given tool to evaluate the gait within trot-ups using different handlers at different vet gates. It was shown before that fatigue induces changes in gait pattern in sound horses

(Riber *et al.*, 2006) and this feature would also need to be addressed throughout a competition. Additionally, if such a technology would only be used for those cases when the lameness is not clear, still this would not overcome the subjectivity of the individual OEV to call for a panel and/or the device.

One of the limitations of this study was the use of a convenience sample, with more than half of the responses being obtained in paper, mostly during international competitions attended by the first author. Although the questionnaires were confidential, they were not anonymous to the first author and this may have introduced some bias. The overrepresentation of Europe and 4-star level veterinarians in the survey reflects the level and venues of the competitions where the questionnaires were collected, and the results might not entirely represent the OEVs class opinions.

5. Conclusions

Gait assessment during endurance competitions is challenging and is not perceived as a straightforward task for the majority of the OEVs. The handler trotting the horse is perceived by OEVs as the most important factor limiting their ability to detect lameness. This suggests that rules for endurance competitions should be more detailed and prescriptive concerning how the trot should be performed, including penalties for non-compliance. At the same time, educational videos showing acceptable and unacceptable trotting procedures could be made available. Additionally, trained horse handlers could be made available to assist OEVs at ride venues to perform the trot-ups. Currently, veterinarians only have visual observation over ~25 s of trotting away and back to assess gait in endurance competitions. The introduction of objective gait evaluation technology, whether sensor or video based, has the potential to improve horse welfare, reduce subjectivity, increase consistency, monitor injuries and may be perceived as fairer by competitors.

Supplementary material

Supplementary material can be found online at https://doi.org/10.3920/CEP180058.

International Survey – Lameness detection by FEI official veterinarians in equine endurance competitions.

References

Adair, S., Baus, M., Belknap, J., Bell, R., Boero, M., Bussy, C., Cardenas, F., Casey, T., Castro, J., Davis, W., Erskine, M., Farr, R., Fischer, T., Forbes, B., Ford, T., Genovese, R., Gottschalk, R., Hoge, M., Honnas, C., Hunter, G., Joyce, J., Kaneps, A., Keegan, K., Kramer, J., Lischer, C., Marshall, J., Oosterlinck, M., Radue, P., Redding, R., Reed, S.K., Rick, M., Santschi, E., Schoonover, M., Schramme, M., Schumacher, J., Stephenson, R., Thaler, R., Vedding Neilsen, J. and Wilson, D.A., 2018. Response to Letter to the Editor: Do we have to redefine lameness in the era of quantitative gait analysis. Equine Veterinary Journal 50: 415-417. https://doi.org/10.1111/evj.12820

Adair, S., Baus, M., Bell, R., Boero, M., Bussy, C., Cardenas, F., Casey, T., Castro, J., Davis, W., Erskine, M., Farr, R., Fischer Jr., A., Forbes, B., Ford, T., Genovese, R., Gottschalk, R., Hoge, M., Honnas, C., Hunter, G., Joyce, J., Kaneps, A., Keegan, K., Kramer, J., Labens, R., Lischer, C., Marshall, J., Oosterlinck, M., Radue, P., Redding, R., Reed, S., Rick, M., Santschi, E., Schoonover, M., Schramme, M., Schumacher, J., Stephenson, R., Thaler, R., Nielsen, J.V. and Wilson, D., 2019. Letter to the editor: a response to 'what is lameness and what (or who) is the gold standard to detect it?' Equine Veterinary Journal 51: 270-272. https://doi.org/10.1111/evj.13043

American Endurance Riding Conference (AERC), 2008a. Endurance rider's handbook. Available at: https://aerc.org/static/AERC_Rider_Handbook.pdf.

American Endurance Riding Conference (AERC), 2008b. Guidelines for judging AERC endurance competitions. Available at: https://aerc.org/static/upload/2009controljudgehb.pdf.

Bathe, A.P., Judy, C.E. and Dyson, S., 2018. Letter to the editor: do we have to redefine lameness in the era of quantitative gait analysis? Equine Veterinary Journal 50(2): 273. https://doi.org/10.1111/evj.12791

Baxter, G.M. and Stashak, T.S., 2011. Examination for lameness: history, visual exam, palpation and manipulation. In: Baxter, G.M. (ed.) Adam's and Stashak's lameness in horses. Wiley-Blackwell, Philadelphia, PA, USA.

Coombs, S.L. and Fisher, R.J., 2012. Endurance riding in 2012: too far too fast? Veterinary Journal 194: 270-271. https://doi.org/10.1016/j.tvjl.2012.10.037

Dalin, G., Magnusson, L.-E. and Thafvelin, B.C., 1985. Retrospective study of hindquarter asymmetry in Standardbred Trotters and its correlation with performance. Equine Veterinary Journal 17: 292-296. https://doi.org/10.1111/j.2042-3306.1985.tb02501.x

Dyson, S., 2014. Recognition of lameness: man versus machine. Veterinary Journal 201: 245-248. https://doi.org/10.1016/j.tvjl.2014.05.018

Dyson, S., 2019. Letter to the editor: continued debate about what constitutes lameness. Equine Veterinary Journal 51: 556-556. https://doi.org/10.1111/evj.13118

Dyson, S. and Greve, L., 2016. Subjective gait assessment of 57 sports horses in normal work: a comparison of the response to flexion tests, movement in hand, on the lunge, and ridden. Journal of Equine Veterinary Science 38: 1-7. https://doi.org/10.1016/j.jevs.2015.12.012

Dyson, S.J., 2009. The clinician's eye view of hindlimb lameness in the horse: technology and cognitive evaluation. Equine Veterinary Journal 41: 99-100. https://doi.org/10.2746/042516409X399963

Fédération Equestre Nationale (FEI), 2013. Endurance pilot global endurance injuries study (GEIS). Fédération Equestre Nationale, Lausanne, Switzerland.

Fédération Equestre Nationale (FEI), 2018a. Endurance rules, updated 9th edition. Fédération Equestre Nationale, Lausanne, Switzerland.

Fédération Equestre Nationale (FEI), 2018b. FEI endurance rules 2018. Notes for guidance. Fédération Equestre Nationale, Lausanne, Switzerland.

Fédération Equestre Nationale (FEI), 2018c. Veterinary regulations, 14th edition. Fédération Equestre Nationale, Lausanne, Switzerland.

Fuller, C.J., Bladon, B.M., Driver, A.J. and Barr, A.R., 2006. The intra- and inter-assessor reliability of measurement of functional outcome by lameness scoring in horses. Veterinary Journal 171: 281-286. https://doi.org/10.1016/j.tvjl.2004.10.012

Greve, L. and Dyson, S., in press. What can we learn from visual and objective assessment of non-lame and lame horses in straight lines, on the lunge and ridden? Equine Veterinary Education. https://doi.org/10.1111/eve.13016

Hewetson, M., Christley, R.M., Hunt, I.D. and Voute, L.C., 2006. Investigations of the reliability of observational gait analysis for the assessment of lameness in horses. Vet Record 158: 852-857. https://doi.org/10.1136/vr.158.25.852

Holbrook, T.C., 2011. The endurance horse. In: Baxter, G.M. (ed.) Adams & Stashak's lameness in horses. Willey-Blackwell, Oxford, UK, pp. 1055-1061.

Ishihara, A., Bertone, A.L. and Rajala-Schultz, P.J., 2005. Association between subjective lameness grade and kinetic gait parameters in horses with experimentally induced forelimb lameness. American Journal of Veterinary Research 66: 1805-1815. https://doi.org/10.2460/ajvr.2005.66.1805

Keegan, K.G., 2007. Evidence-based lameness detection and quantification. Veterinary Clinics of North America: Equine Practice 23: 403-423. https://doi.org/10.1016/j.cveq.2007.04.008

Keegan, K.G., 2011. Objective assessment of lameness. In: Baxter, G.M. (ed.) Adams & Stashak's lameness in horses. Willey-Blackwell, Oxford, UK, pp. 154-164.

Keegan, K.G., Dent, E.V., Wilson, D.A., Janicek, J., Kramer, J., Lacarrubba, A., Walsh, D.M., Cassells, M.W., Esther, T.M., Schiltz, P., Frees, K.E., Wilhite, C.L., Clark, J.M., Pollitt, C.C., Shaw, R. and Norris, T., 2010. Repeatability of subjective evaluation of lameness in horses. Equine Veterinary Journal 42: 92-97. https://doi.org/10.2746/042516409X479568

Keegan, K.G., Kramer, J., Yonezawa, Y., Maki, H., Pai, P.F., Dent, E.V., Kellerman, T.E., Wilson, D.A. and Reed, S.K., 2011. Assessment of repeatability of a wireless, inertial sensor-based lameness evaluation system for horses. American Journal of Veterinary Research 72: 1156-1163. https://doi.org/10.2460/ajvr.72.9.1156

Keegan, K.G., Wilson, D.A., Wilson, D.J., Smith, B., Gaughan, E.M., Pleasant, R.S., Lillich, J.D., Kramer, J., Howard, R.D., Bacon-Miller, C., Davis, E.G. May, K.A., Cheramie, H.S., Valentino, W.L. and Van Harreveld, P.D., 1998. Evaluation of mild lameness in horses trotting on a treadmill by clinicians and interns or residents and correlation of their assessments with kinematic gait analysis. American Journal of Veterinary Research 59: 1370-1377.

Lopes, M.A.F., Eleuterio, A. and Mira, M.C., 2018. Objective detection and quantification of irregular gait with a portable inertial sensor-based system in horses during an endurance race – a preliminary assessment. Journal of Equine Veterinary Science 70: 123-129. https://doi.org/10.1016/j.jevs.2018.08.008

Mackay-Smith, M., Bentham, B., Cohen, M., Nelson, T., Ridgway, K. and Steere, J., 2016. Guidelines for control judges and treatment veterinarians at AERC endurance competitions. Available at: https://aerc.org/static/upload/CJHandbook2016.pdf.

MacKinnon, M.C., Bonder, D., Boston, R.C. and Ross, M.W., 2015. Analysis of stress fractures associated with lameness in Thoroughbred flat racehorses training on different track surfaces undergoing nuclear scintigraphic examination. Equine Veterinary Journal 47: 296-301. https://doi.org/10.1111/evj.12285

Marlin, D. and Williams, J., 2018. Equine endurance race pacing strategy differs between finishers and non-finishers in 120 km single-day races. Comparative Exercise Physiology 14: 11-18. https://doi.org/10.3920/CEP170027

McCracken, M.J., Kramer, J., Keegan, K.G., Lopes, M., Wilson, D.A., Reed, S.K., LaCarrubba, A. and Rasch, M., 2012. Comparison of an inertial sensor system of lameness quantification with subjective lameness evaluation. Equine Veterinary Journal 44: 652-656. https://doi.org/10.1111/j.2042-3306.2012.00571.x

Misheff, M.M., 2011. Lameness in endurance horses. In: Ross, M.W. and Dyson, S.J. (ed.) Diagnosis and management of lameness in the horse. Elsevier Saunders, St. Louis, MO, USA, pp. 1137-1149.

Nagy, A., Dyson, S.J. and Murray, J.K., 2012. A veterinary review of endurance riding as an international competitive sport. Veterinary Journal 194: 288-293. https://doi.org/10.1016/j.tvjl.2012.06.022

Nagy, A., Dyson, S.J. and Murray, J.K., 2013. Riders' prediction of results at Federation Equestre Internationale (FEI) endurance rides and sources of bias in questionnaires completed by riders. Preventive Veterinary Medicine 112: 378-386. https://doi.org/10.1016/j.prevetmed.2013.08.005

Nagy, A., Dyson, S.J. and Murray, J.K., 2017. Veterinary problems of endurance horses in England and Wales. Preventive Veterinary Medicine 140: 45-52. https://doi.org/10.1016/j.prevetmed.2017.02.018

Nagy, A., Murray, J.K. and Dyson, S., 2010. Elimination from elite endurance rides in nine countries: a preliminary study. Equine Veterinary Journal: 637-643. https://doi.org/10.1111/j.2042-3306.2010.00220.x

Nagy, A., Murray, J.K. and Dyson, S.J., 2014a. Descriptive epidemiology and risk factors for eliminations from Federation Equestre Internationale endurance rides due to lameness and metabolic reasons (2008-2011). Equine Veterinary Journal 46: 38-44. https://doi.org/10.1111/evj.12069

Nagy, A., Murray, J.K. and Dyson, S.J., 2014b. Horse-, rider-, venue- and environment-related risk factors for elimination from Federation Equestre Internationale endurance rides due to lameness and metabolic reasons. Equine Veterinary Journal 46: 294-299. https://doi.org/10.1111/evj.12170

Parente, E.J., Russau, A.L. and Birks, E.K., 2002. Effects of mild forelimb lameness on exercise performance. Equine Veterinary Journal 34: 252-256. https://doi.org/10.1111/j.2042-3306.2002.tb05428.x

Parkes, R.S., Weller, R., Groth, A.M., May, S. and Pfau, T., 2009. Evidence of the development of 'domain-restricted' expertise in the recognition of asymmetric motion characteristics of hindlimb lameness in the horse. Equine Veterinary Journal 41: 112-117. https://doi.org/10.2746/042516408X343000

Pleasant, R.S., Baker, G.J., Muhlbauer, M.C., Foreman, J.H. and Boero, M.J., 1992. Stress reactions and stress fractures of the proximal palmar aspect of the third metacarpal bone in horses: 58 cases (1980-1990). Journal of the American Veterinary Medical Association 201: 1918-1923.

Rhodin, M., Egenvall, A., Haubro Andersen, P. and Pfau, T., 2017. Head and pelvic movement asymmetries at trot in riding horses in training and perceived as free from lameness by the owner. PLoS ONE 12: e0176253. https://doi.org/10.1371/journal.pone.0176253

Riber, C., Cuesta, I., Munoz, A., Gata, J., Trigo, P. and Castejon, F.M., 2006. Equine locomotor analysis on vet-gates in endurance events. Equine Veterinary Journal 38: 55-59. https://doi.org/10.1111/j.2042-3306.2006.tb05513.x

Ross, M.V., 2011. Movement. In: Ross, M.V. and Dyson, S.J. (ed.) Diagnosis and management of lameness in the horse. Elsevier Saunders, St. Louis, MO, USA, pp. 64-80.

Serra Braganca, F., Roepstorff, C., Latif, S., Gunst, S., Arpagaus, S., Van Weeren, P.R. and Weishaupt, M.A., 2018. The effect of trotting speed on the kinematics of head, withers and pelvis. In: McKenzie, E.C. (ed.) International Conference on Exercise Physioloy (ICEEP). Comparative Exercise Physiology, Lorne, Australia, pp. S53.

Sloet Van Oldruitenborgh-Oosterbaan, M.M., 2018. Subjective and objective evaluations of horses for a fit-to-compete or unfit-to-compete judgement. 10th International Conference on Equine Exercise Physiology. Comparative Exercise Physiology, Lorne, Australia, p. S49.

Starke, S.D. and Oosterlinck, M., 2019. Reliability of equine visual lameness classification as a function of expertise, lameness severity and rater confidence. Veterinary Record 184: 63-63. https://doi.org/10.1136/vr.105058

Starke, S.D., Raistrick, K.J., May, S.A. and Pfau, T., 2013. The effect of trotting speed on the evaluation of subtle lameness in horses. Veterinary Journal 197: 245-252. https://doi.org/10.1016/j.tvjl.2013.03.006

Tomei, G., Cinti, M.E., Cerratti, D. and Fioravanti, M., 2006. Attention, repetitive works, fatigue and stress. Annali di Igiene: Medicina Preventiva e di Comunità 18: 417-429.

Van Weeren, P.R., Pfau, T., Rhodin, M., Roepstorff, L., Serra Braganca, F. and Weishaupt, M.A., 2017. Do we have to redefine lameness in the era of quantitative gait analysis? Equine Veterinary Journal 49: 567-569. https://doi.org/10.1111/evj.12715

Van Weeren, P.R., Pfau, T., Rhodin, M., Roepstorff, L., Serra Bragança, F. and Weishaupt, M.A., 2018. What is lameness and what (or who) is the gold standard to detect it? Equine Veterinary Journal 50: 549-551. https://doi.org/10.1111/evj.12970

Younes, M., Barrey, E., Cottin, F. and Robert, C., 2016. Elimination in long-distance endurance rides: insights from the analysis of 7,032 starts in 80 to 160 km competitions. Comparative Exercise Physiology 12: 157-167. https://doi.org/10.3920/CEP160022

Supplementary online material of Comparative Exercise Physiology
DOI: https://doi.org/10.3920/CEP180058

Challenges encountered by Federation Equestre Internationale (FEI) veterinarians in gait evaluation during FEI endurance competitions: an international survey

M.C. de Mira, C. Santos, M.A. Lopes and D.J. Marlin

1

INTERNATIONAL SURVEY - Lameness detection by FEI official veterinarians in equine endurance competitions

Assurance of confidentiality: This questionnaire is part of a survey to identify the challenges for horse veterinarians performing lameness evaluation during endurance competitions. The information provided in this questionnaire will remain anonymous. You are being asked to provide your name simply to avoid asking you to fill this questionnaire out twice. Information that can lead to the identification of the participants of this survey will never be disclosed to anyone. Only the authors of this study, Dr. Marco Lopes (University of Missouri, USA) and Dr. Monica Mira (Equimuralha, Portugal) will have access to the filled-out questionnaires.

Definition of lameness detection: in this questionnaire lameness detection means to find if the horse has significant lameness (that justifies elimination from the competition) or not without the need to identify the affected limb(s).

1- Personal information: Name: ____________________________________ Star level: ________
Country(ies) where you currently work as a horse veterinarian: ____________________________________
Year of graduation from veterinary school: ________ Number of years as a horse veterinarian: ________
Number of years evaluating endurance horses during competitions: ________
Number of years as a FEI veterinary official in endurance competitions: ________

2- What is your opinion about lameness detection during endurance competitions?
☐ Any horse veterinarian can perform lameness detection after minimal training.
☐ For an experienced and well trained veterinarian, lameness detection is straightforward work.
☐ Even for an experienced and well trained veterinarian, lameness detection is challenging work.

3- What are your feelings about your ability to detect lameness in horses during an endurance ride?
☐ I am an expert on detecting lameness and do not have any trouble detecting lameness during endurance competitions.
☐ I am a very experienced veterinarian but, many times it is hard for me to define if the horse is lame or not.
☐ I am still learning how to detect lameness in horses and this is why I cannot detect all lameness yet.

4- Based on your personal experience which of the following factors may compromise your ability to detect lameness during an endurance competition?
A- Horse behaving badly during lameness evaluation
☐Yes ☐No
Frequency of this problem: ☐common ☐uncommon ☐very uncommon ☐I have never seen this problem

B- Rider or associate not running the horse well during lameness evaluation
☐Yes ☐No
Frequency of this problem: ☐common ☐uncommon ☐very uncommon ☐I have never seen this problem

C- Inappropriate alley for lameness evaluation (i.e., very irregular, muddy)
☐Yes ☐No
Frequency of this problem: ☐common ☐uncommon ☐very uncommon ☐I have never seen this problem

D- Unsatisfactory lighting of the alley for lameness evaluation
☐Yes ☐No
Frequency of this problem: ☐common ☐uncommon ☐very uncommon ☐I have never seen this problem

E- Heavy rainfall during lameness evaluation
☐Yes ☐No
Frequency of this problem: ☐common ☐uncommon ☐very uncommon ☐I have never seen this problem

1

F- Optic illusions created by pelvic asymmetry (e.g., one side of pelvis higher than the other side)
☐Yes ☐No
Frequency of this problem: ☐common ☐uncommon ☐very uncommon ☐I have never seen this problem

G- Elevated carriage of the tail blocking the view of the pelvis
☐Yes ☐No
Frequency of this problem: ☐common ☐uncommon ☐very uncommon ☐I have never seen this problem

H- Uneven carriage of the tail blocking the view of one side of the pelvis
☐Yes ☐No
Frequency of this problem: ☐common ☐uncommon ☐very uncommon ☐I have never seen this problem

I- Paint horse with asymmetric hair coat color over the pelvis and/or proximal part of the hind limbs
☐Yes ☐No
Frequency of this problem: ☐common ☐uncommon ☐very uncommon ☐I have never seen this problem

J- Horse with spotted hair coat over the pelvis and/or proximal part of the hind limbs (e.g., appaloosa)
☐Yes ☐No
Frequency of this problem: ☐common ☐uncommon ☐very uncommon ☐I have never seen this problem

K- Horse hair coat color that minimally contrasts with the background
☐Yes ☐No
Frequency of this problem: ☐common ☐uncommon ☐very uncommon ☐I have never seen this problem

L- Evaluator's fatigue at the end of the day and/or after examining a large number of horses
☐Yes ☐No
Frequency of this problem: ☐common ☐uncommon ☐very uncommon ☐I have never seen this problem

5- Have you ever been confronted by a rider or an associate (e.g., horse owner, trainer, member of the support team) about your decision to eliminate a horse simply due to lameness?
☐Yes ☐No
Frequency of such a confrontation: ☐common ☐uncommon ☐very uncommon

6- Imagine that you have access to an easy-to-use tool to help you objectively detect and quantify lameness and document the results of the evaluations during an endurance competition without interfering with the pace of your work and without bothering the horses. Would you be interested in using such a tool during an endurance competition?
☐ Yes. Even for lamenesses that I would not have any problem seeing on my own it would be great to use a tool to support and document my findings.
☐ Yes. Only for difficult lamenesses when I am not sure if the horse has significant lameness or not.
☐ No. I do not see any benefit in using a tool to help me detect lameness and document the results of lameness evaluation during endurance competitions.
☐ Other answer (unlimited space for your answer): ____________________________________

7- Additional comments (unlimited space for your answer): ____________________________________

2

CHAPTER III: Objective Detection and Quantification of Irregular Gait with a Portable Inertial Sensor-Based System in Horses during an Endurance Race – A Preliminary Assessment

Marco A.F. Lopes, Ângela Eleutério, Monica C. de Mira (2018). Objective Detection and Quantification of Irregular Gait With a Portable Inertial Sensor-Based System in Horses During an Endurance Race—a Preliminary Assessment. Journal of Equine Veterinary Science. 2018;70:123-9. https://doi.org/10.1016/j.jevs.2018.08.008

Journal of Equine Veterinary Science 70 (2018) 123–129

Contents lists available at ScienceDirect

Journal of Equine Veterinary Science

journal homepage: www.j-evs.com

Original Research

Objective Detection and Quantification of Irregular Gait With a Portable Inertial Sensor-Based System in Horses During an Endurance Race—a Preliminary Assessment

Marco A.F. Lopes [a,*,1], Angela Eleuterio [b], Monica C. Mira [c]

[a] Department of Veterinary Medicine and Surgery, College of Veterinary Medicine, Columbia, MO
[b] Catarina Afonso—Clínica Veterinária, Unipessoal, Lda, Ajustrel, Portugal
[c] AL EQUINE—Innovative Veterinary Services, Evora, Portugal

ARTICLE INFO

Article history:
Received 3 December 2017
Received in revised form
2 August 2018
Accepted 13 August 2018
Available online 27 August 2018

Keywords:
Horse
Lameness
Gait analysis
Sports medicine
Horse welfare

ABSTRACT

Detection of irregular gait (IG) during endurance competitions is based on a brief subjective veterinary examination. Horse elimination due to IG is common and may elicit competitors' complaints. We hypothesized that detection of IG by objectively assessing motion asymmetry can be performed during endurance competitions. The aim of this preliminary study was to start investigating the feasibility of detecting and quantifying IG with a portable inertial sensor-based system (PISBS) during endurance races. Horses participating in two qualifying endurance rides were simultaneously and independently evaluated by Fédération Equestre Internationale veterinarians and the PISBS, and the results were compared after the rides. Asymmetric vertical displacement of the head and/or pelvis measured with the PISBS indicated severity of IG and identified the affected limb(s). Veterinarians and competitors answered questionnaires about the use of the PISBS. The PISBS detected IG in 21/22 horses (48/70 evaluations). Significant disagreement between the PISBS and veterinarians was detected. Disagreement between the PISBS and veterinarians was no longer detected after reducing the sensitivity of the PISBS by reclassifying horses with mild IG as sound. Competitors and veterinarians had favorable impressions about the use of the PISBS but recommended reducing instrumentation time and trotting distance for expediency. Inherent human limitations can explain the lower sensitivity of veterinary evaluation relative to PISBS evaluation. Simple methodological changes are likely to address the issues raised in this study by competitors and Fédération Equestre Internationale veterinarians. Additional studies are needed to complete the feasibility assessment of objective detection and quantification of IG during endurance competitions with a PISBS.

Animal welfare/ethical statement: We hereby state that this study was conducted following the EU Directive 2010/63/EU for animal experiments.
1. No invasive procedure was performed for data collection.
2. Data were collected while horses were been routinely evaluated by veterinarians immediately before, during, and after an endurance race.
3. Competitors and veterinarians voluntarily agreed to participate in the study after being informed about the methods that included confidentiality of all information that could lead to identification of horses and competitors.
4. This research was approved by the Animal Care and Use Committee, University of Missouri, Columbia, MO, USA.

Conflict of interest statement
1. M.L. worked at the University of Missouri where the PISBS used in this study was developed. The University of Missouri owns the patent of this system. M.L. has worked sporadically as a technical consultant for Equinosis LLC, Saint Louis, Missouri, USA, the manufacturer and distributor of the PISBS.
2. Equinosis LLC, Saint Louis, Missouri, USA, the manufacturer and distributor of the PISBS, supported this study by providing equipment.
3. This study was also partially supported by the E. Paige Laurie Endowed Program in Equine Lameness, College of Veterinary Medicine, University of Missouri, Columbia, Missouri, USA.
4. M.M. is an FEI Official Veterinarian but she neither worked as one of the FEI Official Veterinarians in the endurance rides used for this study nor acted as a representative of the FEI while planning the study, collecting, and analyzing data, or preparing this article.

* Corresponding author at: Dr Marco A.F. Lopes, Equine Health & Performance Centre, School of Animal and Veterinary Sciences, University of Adelaide, 1454 Mudla Wirra Rd, Roseworthy, SA 5371, Australia.
E-mail address: maflopes@gmail.com (M.A.F. Lopes).
[1] Current address: Equine Health and Performance Centre, School of Animal and Veterinary Sciences, University of Adelaide, 1454 Mudla Wirra Rd, Roseworthy, SA, 5371, Australia. E-mail: maflopes@gmail.com. Phone: +61 8 8313 9999 Extension 2. Mobile phone: +61 499 354 557.

https://doi.org/10.1016/j.jevs.2018.08.008
0737-0806/© 2018 Elsevier Inc. All rights reserved.

1. Introduction

Detection of irregular gait (IG) (i.e., lameness) is a major part of the veterinary evaluations performed during endurance competitions and is paramount for the well-being of the horses [1–3]. Because most IGs are attributed to musculoskeletal pain, subjecting a horse with IG to intense exercise may aggravate an existing lesion(s) or lead to additional lesions [1,4–6]. Irregular gait may also compromise performance and lead to metabolic derangement [7,8]. Equine endurance, a sport regulated worldwide by the Fédération Equestre Internationale (FEI), is currently the fastest growing equestrian sport. The rapid development of the sport along with a recent report of some types of severe injuries in endurance horses previously thought to occur exclusively in racehorses [4] has raised concerns about the effectiveness of the current methods used to assess the musculoskeletal fitness of horses participating in endurance competitions [9]. Endurance competitions are very long races, and horses can be eliminated by FEI veterinarians for IG per se at any time, even after they cross the finish line [1–3]. In three recent studies with thousands of endurance horse competition starts, elimination due to IG was the most common reason for not completing the competitions (47.3%, 60.9%, and 30.0%, respectively) [10–12]. Elimination due to IG affects the outcome of the races and commonly elicits complaints from endurance competitors [13].

The FEI has been actively modifying its rules for endurance competitions with the aim of addressing concerns with horse welfare associated with IG, ensuring consistency of how the horses are evaluated for IG, and assuring competitors, the media, and the public that veterinary assessment is fair [3]. For example, currently, the ultimate decision about elimination or not of any horse due to IG is made by an independent voting panel consisting of three veterinarians [3]. This approach may work well when the horses have more obvious IG. However, this approach may not be appropriate when the horses have mild to moderate IG because, in these instances, disagreement within the veterinary team and complaints from competitors are more likely.

Due to logistical, financial, and time constraints, examination for IG during endurance competitions is currently limited to inspection of the horse trotted in hand, back and forth, on a 30–50 m straight line [1–3]. Subjective gait evaluation has many well-recognized limitations [3,5,6,12,14–26]. Because detection of IG during endurance competitions is so important for horse welfare and race outcome, and because detection of IG based exclusively on a brief subjective examination has major limitations, the use of an objective gait analysis system to support detection of IG by veterinarians would likely benefit all segments of the equine endurance industry.

Until recently the use of objective gait analysis for detection of IG outside the laboratory was impractical and cost-prohibitive because it required expensive equipment and software, complicated horse instrumentation, and labor-intensive data collection and analysis [27,28]. This paradigm has been changed by new technologies that have made objective detection and quantification of IG in horses exercising outdoors relatively easy, fast, and inexpensive [28].

The main objective of this preliminary study was to start assessing the feasibility of detecting and quantifying IG during endurance competitions with a portable inertial sensor-based system (PISBS). We hypothesized that the portable PISBS for detection and quantification of IG in horses can be used during endurance competitions without causing significant disruption of the usual activities performed by the FEI veterinarians.

2. Materials and Methods

2.1. Subjects

All competitors of two qualifying endurance rides (40 km, with a stop at 20 km; and 80 km, with stops at 40 and 60 km) simultaneously held in Fronteira, Portugal, were invited to participate in a study where all gait evaluations (before, during, and after the competition) were simultaneously performed using two approaches: veterinary evaluation based on FEI rules and evaluation with a portable PISBS (Lameness Locator, Equinosis LLC, Saint Louis, MO, USA). The endurance rides were not part of any championship and prizes worth less than US$100 were given to the winners. Horse speed was controlled (range, 12–16 Km/h). The maximum heart rate at the vet gates was set at 64 beats/minute (within 20 minutes after arrival and, for the final evaluation, within 30 minutes after crossing the finish line). The cardiac recovery index (CRI) was not measured in the qualifying rides used for this study. Competitors and veterinarians voluntarily agreed to participate in the study after being informed about the methods that included confidentiality of all information that could lead to identification of horses, competitors, or veterinarians.

2.2. Simultaneous Subjective and Objective Lameness Evaluations

Although the horses in the vet gate were being examined by FEI veterinarians before, during, and after the rides, they were non-invasively fitted with a set of the three small inertial sensing devices, each device weighing about 30 g: two devices (each one containing an uniaxial accelerometer to measure vertical acceleration of the head and pelvis) were attached to the dorsal midline, on the poll, and over the sacral tuberosities; the third device (containing an uniplanar gyroscope to measure flexion and extension of the digit on the sagittal plane) was attached to the dorsum of the pastern of the right forelimb [28,29]. Instrumentation was always performed immediately after the heart rate had been assessed by the FEI official veterinarian due to concerns that fitting the horses with the sensors could cause an increase in heart rate. Time required for horse instrumentation was not recorded. In accordance with FEI rules [1–3], all horses were evaluated for IG immediately before the ride and after completing each segment of the ride. For gait evaluation, each horse was trotted in hand back and forth in a straight line on a very even and hard dirt alley about 35 m long by the rider or an associate. Whenever the veterinarian observed consistent IG or even suspected that the horse was lame, a final decision on elimination from the ride was made by a panel of three FEI veterinarians voting independently. The only modifications to the standard FEI gait examination protocol for endurance competitions [3] were horse instrumentation with the three lightweight measuring devices and the requirement to trot the horses twice the usual distance (back and forth twice instead of just once). The last change was made to maximize the chance of collecting data from at least 25 strides deemed by the PISBS as acceptable strides as recommended by the manufacturer of the PISBS [28]. Three units of the PISBS were used; each unit composed of a tablet computer and two sets of wireless measuring devices. A veterinarian familiar with the PISBS operated each unit and was assisted by two individuals responsible for horse instrumentation. After the endurance rides, information about the findings of all veterinary evaluations was obtained from the organizers of the rides.

2.3. Analysis of PISBS Raw Data

Each data set collected with the PISBS typically contained four sequences of strides at the trot. Data were analyzed immediately after collection using PISBS software. Data analysis started with selection of good strides at the trot from each one of the four stride sequences based on the following criteria: (1) frequency of head and pelvic vertical oscillation had to be equal to twice the frequency of right forelimb motion; (2) each stride sequences had to be at least six strides long; (3) stride rate had to be between 90% and 110% of the median stride rate; and (4) for each stride, the square root of the sum of $Mindiff_{head}^2$ and $Maxdiff_{head}^2$ had to be within the 95% confidence interval of the mean (mean ± 2 standard deviations) for all strides sequences combined. Criteria recommended by the manufacturer were used to interpret the results of the PISBS, that is, horses were considered to have front limb lameness when the absolute values of the mean of $Mindiff_{head}$ and/or $Maxdiff_{head}$ were above threshold (6 mm) and the standard deviation of the mean(s) ($Mindiff_{head}$ and/or $Maxdiff_{head}$) was smaller than the respective mean; horses were considered to have hind limb lameness when the absolute values of the mean of $Mindiff_{pelvis}$ and/or $Maxdiff_{pelvis}$ were above threshold (3 mm) and the standard deviation of the mean(s) ($Mindiff_{pelvis}$ and/or $Maxdiff_{pelvis}$) was smaller than the respective mean. For comparison with the evaluations performed by FEI veterinarians, front limb and hind limb IG detected with the PISBS were graded as mild when the absolute value of Mindiff and/or Maxdiff was above threshold but no greater than 2× threshold and moderate to severe lameness when the absolute value of Mindiff and/or Maxdiff was larger than 2× threshold. To avoid biasing the FEI veterinarians and instigating any criticism of the veterinary evaluation, competitors and official veterinarians did not have access to the results of PISBS evaluations on the day of the rides.

2.4. Assessment of Competitors' and FEI Veterinarians' Impressions About the PISBS

At the end of the rides, all competitors (riders or owners) and official veterinarians were asked to fill a quantitative questionnaire about the use of the PISBS during the endurance competition. One week after the rides, the results of the PISBS evaluations were sent to each FEI veterinarian (all the PISBS evaluations were matched with the veterinary evaluations without identifying horses, riders, or veterinarians) and to each one of the competitors (only the results of the PISBS evaluations of their respective horses). Fédération Equestre Internationale official veterinarians were then interviewed by phone and/or e-mail.

2.5. Classification of Horses as Having Acceptable Gait or Not

For each trial of each horse, IGs were considered to be present or absent based on three approaches: (1) veterinary evaluation—-when the FEI veterinarian considered the horse to be lame or suspected that the horse was lame irrespective of whether this led to elimination from the ride (grade C) or not (grade B) [3]; (2) standard PISBS evaluation—when the PISBS detected IG of any severity; and (3) low sensitivity PISBS evaluation—when the PISBS detected moderate to severe IG.

2.6. Statistical Analysis

Data from the following types of trials were not included in the final analysis: repeated evaluations at the same vet gate (reinspections and evaluations performed by a panel of veterinarians after the horse was considered to be lame or suspected to be lame by the primary veterinarian) and trials for which the results of the subjective gait evaluation were not recorded. To investigate if collecting shorter data sets (i.e., running the horses only once back and forth for 35 m) with the PISBS would have compromised the sensitivity of detection of IG; each data set with four trot sequences and at least 25 strides selected by the PISBS was shortened by removing the last two trot sequences and analyzed as performed for the full data sets.

For each vet check with at least 15 horses evaluated, the McNemar's test and odds ratios were used to make the following comparisons between the different approaches for lameness detection: (1) veterinary inspection versus standard PISBS evaluation; (2) veterinary inspection versus the low sensitivity PISBS evaluation; and (3) analysis of each PISBS data set for the full four trot sequences versus analysis of just the first two trot sequences of each data set. Final placements in the rides of horses evaluated and not evaluated with the PISBS were compared using the Mann–Whitney test. Among the horses eliminated from the 40 km ride, the proportions of horses evaluated and not evaluated with the PISBS were calculated and compared using the Fisher's exact test. Statistical analysis was performed with commercial software (Microsoft Excel, Microsoft Corporation, Washington, USA and SAS, SAS Institute Inc., North Carolina, USA).

3. Results

3.1. Horses and PISBS Evaluations

Twenty-nine horses participated in the rides, 22 in the 40 km ride and 7 in the 80 km ride. Horses ranged in age from 5 to 11 years (40 km, median, 6; 80 km, median 7). Twenty-two competitors (78.6%) (40 km class, n = 15 [71.4%]; 80 km class, n = 7 [100%]) agreed to have their horses evaluated with the PISBS. Throughout the day, there was no rain and the excellent condition of the ground for gait evaluation was well preserved. Data from 7 of the 77 evaluations with the PISBS were excluded from the final analysis: six reevaluations when IGs were detected in the first inspection, and one trial for which the gait grade was not recorded by the veterinarian because the horse was eliminated due to metabolic derangement. The PISBS detected IG in 48/70 of the evaluations (68.6%). Of the 48 evaluations where IGs were detected by the PISBS, a single limb was affected in 32 evaluations (66.7%), and two limbs or more were affected in 16 evaluations (33.3%). Of the 48 evaluations where IGs were detected by the PISBS, front limb IGs were detected in 13 evaluations (27.1%), hind limb IGs were detected in 20 evaluations (41.7%), and simultaneous front and hind limb IGs were detected in 15 evaluations (31.3%). Of the 48 evaluations where IGs were detected by the PISBS, front limb IGs were the only IG or the most severe IG detected in 18 evaluations (37.5%), hind limb IGs were the only IG or the most severe IG detected in 22 evaluations (45.8%), and front and hind limb IGs with the same severity were detected in eight evaluations (16.7%). Before the start of the rides, the PISBS indicated that 12/22 horses (54.5%) had IG and 7/22 horses (31.8%) had moderate to severe IG. The PISBS did not detect IG at any time point in only one of the 22 horses (4.5%). Four patterns of IG were detected in the other 21 horses (see supplementary material): (1) IG switching from one limb to the other or going away during subsequent evaluations, 8/22 horses (38.1%); (2) IG affecting the same limb(s) since the beginning of the race, 6/22 horses (28.6%); (3) IGs that appeared during the competition and were detected in the last two or three evaluations, 4/22 horses (19.0%); and (4) IG detected only after the competition, 3/22 horses (14.3%).

3.2. Outcomes of the Rides

Two of the 22 horses were eliminated from the 40 km race, one due to metabolic derangement (evaluated with the PISBS) and one due to IG (not evaluated with the PISBS). Four of the seven horses competing in the 80 km race were eliminated, three due to IG and 1 due to metabolic derangement. Among the horses eliminated from the 40 km ride, there was no difference in the proportion of horses evaluated (1/15) or not evaluated (1/7) with the PISBS ($P = .455$). Among the 40 km horses, those evaluated with the PISBS had better final classification in the race (first to 16th position; median, 7.5th) than the horses not evaluated with the PISBS (12th to 20th position; median 17.5th) ($P = .003$).

3.3. Comparison Between Veterinary Evaluations and PISBS Evaluations

There was significant disagreement between veterinary evaluation and standard PISBS evaluation (i.e., when all IGs detected by the PISBS were considered relevant) (Fig. 1, Table 1). However, no significant disagreement was detected between veterinary evaluation and low sensitivity PISBS evaluation (i.e., when horses with mild IG detected by the PISBS were reclassified as having acceptable gait) (Fig. 1, Table 1).

3.4. Practical Aspects of the PISBS Evaluation

Although attachment of the PISBS devices to the horse was completed in a few minutes, eight of 17 competitors who filled the questionnaire and 1 of 4 FEI veterinarians who officiated at the rides still considered that ideally this procedure needs to be performed even faster. When asked about the use of the PISBS in higher level competitions, all 16 competitors who answered this part of the questionnaire and all veterinarians were positive about it. Five of 17 competitors complained about the longer than usual trot for lameness evaluation. None of the veterinarians thought that this research affected their ability to detect IG.

Since multiple evaluations were performed in 22 horses and three units of the PISBS were used in this study, it was impossible to perform all evaluations for each horse using the same computer. Because there is no practical way to transfer data between the PISBS units, during the qualifying rides, it was impossible to generate reports comparing the results of sequential evaluations for each horse.

3.5. Comparison Between the Full PISBS Data Sets and the First Two Trots of the PISBS Data Sets

After removal of outlier strides, only 55 of the 70 PISBS data sets (78.6%) had the minimum number of strides (25) and could be split in half. The reasons for the PISBS to select fewer than 25 strides for analysis in 15 trials (21.4%) were (1) in 11 trials the horse had very inconsistent trot during part of data collection which resulted in large proportions of strides not fulfilling PISBS criteria and being excluded from the analysis by PISBS software; (2) the operator of the PISBS failed to collect the first sequence of the four sequences of strides in three trials; and (3) the operator of the PISBS

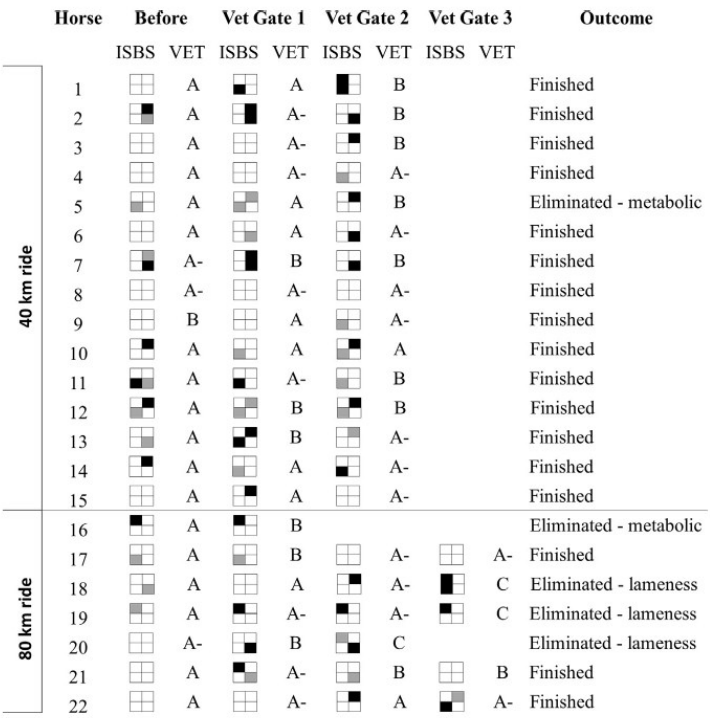


Fig. 1. Lameness detection in 22 horses (70 evaluations) during endurance competitions using two approaches: PISBS lameness——evaluation with a PISBS (boxes with four quadrants, one for each limb [thoracic limbs above; right limbs on the right side]; gray filling = mild lameness [Mindiff and/or Maxdiff > threshold and <2× threshold], black filling = moderate/severe lameness [Mindiff and/or Madiff ≥ 2× threshold]); VET—examination by FEI veterinarians grading gait as "A" (normal), "A-" (mildly abnormal), "B" (no consistently observable gait abnormality), or "C" (consistently observable gait abnormality found by at least 2 among the 3 veterinarians of the panel resulting in elimination due to lameness). Time points: before the ride; after 20 km (Vet Gate 1) and 40 km (Vet Gate 2) for the 40-km ride; after 40 km (Vet Gate 1), 60 km (Vet Gate 2), and 80 km (Vet Gate 3) for the 80-km ride. PISBS, portable inertial sensor-based system; FEI, Fédération Equestre Internationale.

Table 1
Comparisons between lameness evaluations (grade 0 = sound; grade 1 = lame) performed simultaneously by FEI veterinarians and by a PISBS and in equine endurance competitions.

Vet. Evaluation Compared To	Time Point	n	*P*-value (McNemar's)	Odds Ratio	Odds Ratio 95% CI	
					Lower	Upper
Standard ISBS evaluation	Before the race	22	0.003	12.0	1.8	513
	Vet check 1	22	0.002	13.9	2.9	Infinite
	Vet check 2	21[a]	0.004	12.5	2.5	Infinite
Lower sensitivity ISBS evaluation	Before the race	22	0.070	7.0	0.8	56.9
	Vet check 1	22	0.289	3.0	0.6	14.9
	Vet check 2	21[a]	0.180	3.5	0.7	16.8

For the first 3 comparisons, the standard PISBS approach was used (i.e., all lameness detected by the PISBS were deemed significant).
For the last 3 comparisons, the sensitivity of the PISBS was reduced by considering mild lameness not significant.
Odds ratio = odds of detecting lameness with the PISBS/odds of detecting lameness by veterinary evaluation.
[a] One horse was eliminated from the race at vet check 1.

inadvertently finalized data collection before the horse had completed the four trot sequences in one trial. Significant disagreement between horse classification as sound or lame based on the whole PISBS data sets or on the first two trot sequences of each data set was not detected at any time point (Table 2).

4. Discussion

In this preliminary study, it was possible to perform lameness evaluation with a portable PISBS in horses during endurance race without affecting the pace of the competition. The simplicity of data collection and analysis with the PISBS used in this study contrasts with the laborious and time-consuming objective gait analysis with traditional kinematics used in a previous study during an endurance competition [27]. Both at the end of the race and after being informed about the results of the PISBS evaluations, veterinarians and competitors were positive about the use of the PISBS to assist veterinarians in the detection, quantification, and documentation of IG during endurance competitions.

It is important to emphasize that to investigate the accuracy of lameness detection with the PISBS was not a goal of this study. To assess the accuracy of the PISBS, it would have been necessary to use a reference standard [30] such as detailed inspection of videos (at normal speed and in slow motion) of the horses trotting performed by a group of experienced veterinarians. In light of the major limitations of detection of IG based exclusively on a brief subjective evaluation [14–25], it would not be logical to consider the findings of the FEI official veterinarians as the gold standard for lameness detection. Detection of IG based exclusively on subjective evaluation is inherently compromised by human limitations, such as insufficient spatial and temporal resolutions of the human vision [20], incomplete and transient memory of what humans can see [21], and evaluators' bias [22]. Detection of IG based exclusively on subjective evaluation can also be affected by horse speed [24], elevated tail carriage blocking the veterinarian's view of the pelvis when the horse is trotting away from the veterinarian [5,25], optic illusions created by pelvic asymmetry or uneven hair coat color on the pelvis [25], and level of training of the veterinarian [26]. Reliability of subjective detection of IG is further compromised during endurance competitions by time constraints, which unavoidably restricts gait evaluation to a brief inspection and superficial palpation at rest followed by a short inspection of the horse trotting in a 30–50 m straight line away and toward the veterinarian once [3]. This approach for detection of IG during endurance competitions defies the opinion of experts, who recommend a much more detailed work up, which includes inspection of the horse at the walk, inspection of the horse from multiple angles (i.e., from both sides, from behind, and from the front), inspection of the horse moving in circles, inspection of the horse moving on both soft and hard surfaces, evaluation of the response to flexion tests, and detailed palpation of the limbs and torso [6,23].

In accordance with FEI rules [3] and to not introduce bias, in this study, the official veterinarians did not have access to the results of the PISBS evaluation during the rides. Evaluator bias has been shown to affect lameness detection based on veterinary evaluation [22]. However, the PISBS was never meant to be used by itself. Indeed the veterinarian using the PISBS should perform routine lameness evaluation, while data are collected and the findings of the PISBS should be interpreted in light of the veterinarian's own observations [28,29].

Instrumentation with the sensors of the PISBS immediately before the trot, as performed in this study, could potentially affect heart rate after the trot especially in fractious horses. This hypothesis was not tested in this study because the CRI was not measured in the qualifying endurance rides included in this study. If it is demonstrated that the CRI can be affected by instrumentation with the sensors of the PISBS immediately before the trot, as performed in this study, this issue could be solved by instrumenting the horses before they enter the vet check.

The trot alleys were in excellent condition throughout the ride used in this study. Despite the lack of published studies, it is reasonable to expect that the quality of the ground where the horses are trotted can affect both the objective (with the PISBS) and

Table 2
Comparisons between lameness evaluations (grade 0 = sound; grade 1 = lame) using an PISBS with analysis of data sets in full (four trot sequences from horses trotted back and forth twice for 35 m) or just the first two trot sequences of each data set.

Time Point	n	*P*-value (McNemar's)	Odds Ratio	Odds Ratio 95% CI	
				Lower	Upper
Before the race	17	1.000	0.5	0.01	9.6
Vet check 1	17	0.500	2.4	0.29	Infinite
Vet check 2	16[a]	1.000	1.0	0.05	Infinite

Data were collected from 22 horses participating in endurance competitions, but five original data sets had fewer than 4 trot sequences and/or fewer than 25 strides and were not included in the comparisons.
Odds ratio = odds of detecting lameness using the first half of the data set/odds of detecting lameness using the whole data set.
[a] One horse was eliminated from the race at vet check.

the subjective components of gait evaluation. Although no studies assessing the effect of an irregular ground on the performance of the PISBS have been published yet, the authors' personal experience using the PISBS for lameness evaluations is that an irregular ground can increase variability of the measurements and negatively affect the performance of the PISBS. Because the PISBS uses the average and the standard deviation of the differences in vertical acceleration of the head and pelvis associated with the load on each limb, very irregular ground may even put the PISBS in disadvantage relative to subjective evaluation. Just a couple of very convincing asymmetric movements of the head or pelvis may be enough for the veterinarian to correctly see an IG not consistently expressed (i.e., with a very large standard deviation relative to the mean) and therefore undetectable by the PISBS. Thus, to obtain reliable results with a tool like the PISBS, it is essential to have the ground for gait evaluation in good condition at all times. According to the FEI rules [3], "a flat firm surface" is mandatory for gait evaluation during endurance competitions, but this rule may be very difficult or even impossible to enforce. A muddy but relatively flat ground would likely not be a problem for the evaluation with the PISBS. However a very irregular ground, which could also compromise the visual gait assessment performed by the veterinarians, would likely add too much variability to the vertical motion of the head and pelvis, which could compromise the performance of the PISBS. Thus, before adopting the PISBS for gait evaluation at endurance competitions, it would be important to test the PISBS in horses trotting on more irregular surfaces for gait evaluation at FEI endurance rides.

After more studies are conducted, it will be up to the FEI or any other governing body such as the Australian Endurance Riders Association or American Endurance Ride Conference, the organizers of the competitions and the veterinarians to define how the PISBS used in this study or any other tool for objective detection and quantification of lameness could be used in endurance competitions. There are certainly several options for the use of this kind of tool to help veterinarians objectively assess gait and define which horses are fit to continue in the endurance ride such as the following: (1) The PISBS could be used exclusively in contentious situations when the rider/owner disagrees that the horse should be eliminated from the competition due to IG; (2) The PISBS could be used for all evaluations but simply to support the subjective assessment performed by the veterinarian and the horse would only be considered to be lame if both the veterinarian could see the IG and the PISBS could detect moderate to severe lameness; (3) The PISBS could be used exclusively as a tool to assess the progression of IG during the competition and significant worsening of IG affecting the same limb detected by the PISBS at a subsequent vet check could be set as a criterion for elimination from the competition; (4) The PISBS could be used fully and IG above a preestablished threshold, regardless if the veterinarian could see it or not, would be grounds for elimination from the competition. It is likely possible to combine some of these options or even to envision other ways of using the PISBS in endurance competitions. Regardless of the approach used, implementation of the use of this kind of tool for detection of IG during endurance competitions would certainly require training for veterinarians and their assistants, and education of competitors and horse owners.

The disagreement between the PISBS results and the veterinary evaluations observed in the present study was expected because similar disagreement has been found in previous studies comparing subjective to objective lameness evaluation [16–19]. This disagreement may be attributed to human limitations to visually assess motion [20–22], the very brief subjective evaluation performed during endurance competitions [3] and the higher sensitivity of the PISBS when compared to subjective evaluation [19]. The fact that significant disagreement between the veterinary evaluation and the PISBS evaluation was only found when mild IGs detected with the PISBS were considered relevant is evidence of the higher sensitivity of the PISBS.

Additional studies are needed to assess the clinical significance of mild lameness detected by the PISBS. Meanwhile, it would be unreasonable to disqualify horses from endurance competitions solely because, according to the PISBS, those horses have mild IG [28,29]. If this had been done in the rides used for the present study, all but one horse would have been eliminated due to IG and 12 of the 22 horses (54.5%) would not have been allowed even to start in the rides. Although it is likely appropriate to eliminate the horses from the rides found by the PISBS to have moderate to severe IG, further studies are necessary to determine the clinical significance (i.e., the effect on horses' well-being) of IG of different severities and to define the ideal cutoff value for disqualification from endurance competitions. The clinical significance of other variables that can also be easily monitored with the PISBS such as the stage of the ride where IGs are first detected, the limb(s) affected (i.e., lameness is affecting the same limb at several vet checks or lameness is shifting from one limb to the other), and variation of IG severity overtime should also be investigated (i.e., lameness is still the same, is improving, or is worsening).

None of the FEI veterinarians thought that their performances were affected by this study, but it is impossible to account for subconscious bias. Despite the assurances of confidentiality, the veterinarians may have been concerned about having their performances compared to the PISBS. This may have caused the veterinarians to be more likely to either declare very mild IG or to only declare very obvious IG. In addition, the low standard of the rides and the insignificant prize money may have made the veterinarians less strict about eliminating horses due to IG. The effect of bias on the performance of veterinarians conducting lameness evaluations in horses has been clearly demonstrated in a previous study [22].

Although the high incidence of shifting IG from one limb to another at different vet checks has not been documented before, this finding may be consistent with the occurrence of transient lameness previously reported in endurance horses [13]. Several traumatic episodes may occur during an endurance competition and it is possible that the most recent trauma before each vet check could have caused a different lameness from one previously identified. Several factors may explain why IG alternately affecting different limbs had not been reported before; according to FEI rules, the official veterinarians indicate if the horse has IG or not but do not identify the affected limb(s) [3]; the sensitivity of traditional gait evaluation is low [16–19]; typically, during an endurance competition, each horse is not evaluated by the same veterinarian at all vet checks. The clinical significance of transient and intermittent IG during an endurance competition requires further investigation.

As alluded to by competitors and veterinarians, the time required to attach measuring devices and collect data with the PISBS needs to be extremely short to prevent delays during a competition. This issue can be addressed by simple modifications in the methods used in this study: (1) by using device attachments that are even easier and quicker to apply and remove than those recommended by the manufacturer [29]; (2) by providing more training for veterinarians and assistants in the use of PISBS; (3) by starting instrumentation of the next horse while the preceding horse is being evaluated; and (4) by trotting horses back and forth just once for each evaluation. In this study, all horses were trotted for twice the distance typically used in endurance competitions only because the manufacturer of the PISBS recommends collecting data from at least 25 strides for each evaluation [28,29]. If the alleys' length had been closer to the maximal length recommended by the FEI (50 m) [3], it would have been more likely to obtain 25 strides after removal of outlier strides by running the horse back and forth just once. In the present study, it was also demonstrated that, even when the alleys were shorter than 50 m, it was possible to collect enough strides when the horses did not have a very inconsistent trot or the operators of the PISBS did not make any mistake. Because an irregular trot can compromise detection of IG by veterinary inspection, the FEI rules for veterinary evaluation during endurance competitions already indicate that, when the horse does not trot well, the veterinarians can request another trot [3]. Gait evaluation with the PISBS performed simultaneously with traditional gait evaluation during endurance rides could also provide unbiased assessment of the quality of the trot to determine when it is necessary to request a second trot. Additional research is necessary to define appropriate guidelines before the PISBS can be used for this application.

Modifications to hardware and software of the PISBS that were used in this study are also desirable if this PISBS is to be adopted as a tool to help veterinarians detect and quantify IG during endurance competitions: (1) To minimize the time required for horse instrumentation even further, it would be desirable to reduce the number of measuring devices of the PISBS. Since there is evidence that pelvic rotation on the transverse plane can be used to detect the phases of the steps of the hind limbs [31], it may be possible to use pelvic rotation to determine the stride frequency and to infer the phases of the steps of all limbs at any time when the horse is trotting. Therefore, inclusion of a uniplanar gyroscope to assess pelvic rotation on the transverse plane in the inertial sensing device attached to the pelvis combined with software modifications would likely make possible to eliminate the device attached to the right front limb. (2) To minimize even further the time required for data analysis, it would be better to use faster computers and/or to modify PISBS software to perform data analysis simultaneously with data collection. (3) To enable on site comparisons of results, a practical approach for continuous sharing of data between multiple units of the PISBS has to be developed.

Potential benefits of the use of the portable PISBS in endurance competitions and good acceptance by competitors and veterinarians to this new approach for detection of IG have been demonstrated in this preliminary study. The main limitations of this study were (1) the PISBS was evaluated only in two qualifying rides with limited speed; (2) only four FEI official veterinarians participated in the study; (3) only a small number of horses were included in the study; and (4) the level of the endurance ride used for the study was low. Additional experiments with more horses, more rides, more FEI veterinarians, and higher level endurance rides (where the CRI is measured), and with the official veterinarians using the PISBS as it is meant to be used (as a tool to support veterinary decisions instead of a tool used independently from the veterinary evaluations as done in this study) are needed to better investigate the benefits of the PISBS and to complete the feasibility assessment of objective detection and quantification of IG during endurance competitions with a PISBS.

Acknowledgments

The authors acknowledge the support from the following institutions: E. Paige Laurie Endowed Program in Equine Lameness, College of Veterinary Medicine, University of Missouri, Columbia, Missouri, USA; Equimuralha, Évora, Portugal; Alltech Portugal, Sintra, Portugal; Pavo, Lisboa, Portugal.

The authors also thank Mr. António Gomes, Dr. João Candeias, Dr. Miguel Minas, Dr. Nuno Prates, Dr. Tomé Fino, Ms. Ana Pinto, Ms. Cláudia Biller, Ms. Cláudia Ribeiro, Ms. Débora Martins, Ms. Joana Rua, Ms. Joanne McMurray, Ms. Liliane Damásio, Ms. Mafalda Pardal, Dr. Susana Monteiro, and Ms. Marta Tobar for their help with data collection and/or organization of this study.

Supplementary Data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jevs.2018.08.008.

References

[1] Holbrook TC. The endurance horse. In: Baxter GM, editor. Adams & Stashak's lameness in horses. 6th ed. Oxford: Willey-Blackwell; 2011. p. 1055–61.
[2] Misheff MM. Lameness in endurance horses. In: Ross MW, Dyson SJ, editors. Diagnosis and management of lameness in the horse. 2nd ed. St. Louis: Elsevier Saunders; 2011. p. 1137–49.
[3] FEI. Updated 9th edition, effective 1 January 2018. https://inside.fei.org/sites/default/files/Endurance%20Rules_2018_clean.pdf. [Accessed 3 February 2018].
[4] Misheff MM, Alexander GR, Hirst GR. Management of fractures in endurance horses. Equine Vet Educ 2010;22:623–30.
[5] Ross MV. Movement. In: Ross MV, Dyson SJ, editors. Diagnosis and management of lameness in the horse. 2nd ed. St. Louis: Elsevier Saunders; 2011. p. 64–80.
[6] Baxter GM, Stashak TS. History, visual exam, palpation, and manipulation. In: Baxter GM, Stashak TS, editors. Adams and Stashak's lameness in horses. 6th edn. Oxford: Willey-Blackwell; 2011. p. 109–50.
[7] Morris EA, Seeherman HJ. Clinical evaluation of poor performance in the racehorse: the results of 275 evaluations. Equine Vet J 1991;23:169–74.
[8] Parente EJ, Russau AL, Birks EK. Effects of mild forelimb lameness on exercise performance. Equine Vet J Suppl 2002;34:252–6.
[9] Coombs SL, Fisher RJ. Endurance riding in 2012: too far too fast? Vet J 2012;194:270–1.
[10] Fielding CL, Meier CA, Balch OK, Kass PH. Risk factors for the elimination of endurance horses from competition. J Am Vet Med Assoc 2011;239:493–8.
[11] Nagy A, Murray JK, Dyson S. Elimination from elite endurance rides in nine countries: a preliminary study. Equine Vet J Suppl 2010;38:637–43.
[12] Nagy A, Murray JK, Dyson SJ. Descriptive epidemiology and risk factors for eliminations from Federation Equestre Internationale endurance rides due to lameness and metabolic reasons (2008-2011). Equine Vet J 2014;46:38–44.
[13] Nagy A, Dyson SJ, Murray JK. A veterinary review of endurance riding as an international competitive sport. Vet J 2012;194:288–93.
[14] Fuller CJ, Bladon BM, Driver AJ, Barr AR. The intra- and inter-assessor reliability of measurement of functional outcome by lameness scoring in horses. Vet J 2006;171:281–6.
[15] Hewetson M, Christley RM, Hunt ID, Voute LC. Investigations of the reliability of observational gait analysis for the assessment of lameness in horses. Vet Rec 2006;158:852–7.
[16] Ishihara A, Bertone AL, Rajala-Schultz PJ. Association between subjective lameness grade and kinetic gait parameters in horses with experimentally induced forelimb lameness. Am J Vet Res 2005;66:1805–15.
[17] Keegan KG, Dent EV, Wilson DA, et al. Repeatability of subjective evaluation of lameness in horses. Equine Vet J 2010;42:92–7.
[18] Keegan KG, Wilson DA, Wilson DJ, et al. Evaluation of mild lameness in horses trotting on a treadmill by clinicians and interns or residents and correlation of their assessments with kinematic gait analysis. Am J Vet Res 1998;59:1370–7.
[19] McCracken MJ, Kramer J, Keegan KG, et al. Comparison of an inertial sensor system of lameness quantification with subjective lameness evaluation. Equine Vet J 2012;44:652–6.
[20] Sweet AL. Temporal discrimination by the human eye. Am J Psychol 1953;66:185–98.
[21] Todd JJ, Marois R. Capacity limit of visual short-term memory in human posterior parietal cortex. Nature 2004;428:751–4.
[22] Arkell M, Archer RM, Guitian FJ, May SA. Evidence of bias affecting the interpretation of the results of local anaesthetic nerve blocks when assessing lameness in horses. Vet Rec 2006;159:346–9.
[23] Dyson S. Recognition of lameness: man versus machine. Vet J 2014;201:245–8.
[24] Starke SD, Raistrick KJ, May SA, Pfau T. The effect of trotting speed on the evaluation of subtle lameness in horses. Vet J 2013;197:245–52.
[25] Dyson SJ. The clinician's eye view of hindlimb lameness in the horse: technology and cognitive evaluation. Equine Vet J 2009;41:99–100.
[26] Parkes RS, Weller R, Groth AM, May S, Pfau T. Evidence of the development of 'domain-restricted' expertise in the recognition of asymmetric motion characteristics of hindlimb lameness in the horse. Equine Vet J 2009;41:112–7.
[27] Riber C, Cuesta I, Munoz A, et al. Equine locomotor analysis on vet-gates in endurance events. Equine Vet J Suppl 2006;36:55–9.
[28] Keegan KG. Objective assessment of lameness. In: Baxter GM, editor. Adams & Stashak's lameness in horses. 6th edn. Oxford: Willey-Blackwell; 2011. p. 154–64.
[29] Equinosis. Lameness locator TM training manual. http://equinosis.com/wp-content/uploads/2014/07/LL-User-Manual.pdf. [Accessed 6 September 2014].
[30] Manchikanti L, Derby R, Wolfer L, et al. Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: part 5. Diagnostic accuracy studies. Pain Physician 2009;12:517–40.
[31] Starke SD, Witte TH, May SA, Pfau T. Accuracy and precision of hind limb foot contact timings of horses determined using a pelvis-mounted inertial measurement unit. J Biomech 2012;45:1522–8.

CHAPTER IV: Do Waiting Times in Endurance Vet Gates Affect the Cardiac Recovery Index?

Mónica C. de Mira, Jane Williams, Rute Santos, Patrícia Rodrigues, Beatriz Arroja, David Marlin (2020). Comparative Exercise Physiology. 2020:1-8. https://doi.org/10.3920/CEP190081

Do waiting times in endurance vet gates affect the cardiac recovery index?

M.C. de Mira[1,2*], J. Williams[3], R.G. dos Santos[4], P. Rodrigues[5], B. Arroja[6] and D. Marlin[7]

[1]MED – Mediterranean Institute for Agriculture, Environment and Development, Institute for Advanced Studies and Research, Universidade de Évora, Pólo da Mitra, Ap. 94, 7006-554 Évora, Portugal; [2]Faculdade de Medicina Veterinária, Universidade Lusófona de Humanidades e Tecnologia, Campo Grande 376, 1749-024 Lisboa, Portugal; [3]Hartpury University, GL19 3BE, Gloucester, United Kingdom; [4]VALORIZA – Research Centre for Endogenous Resource Valorization, Campus Politécnico 10, 7300-555 Portalegre, Portugal; [5]Departamento de Medicina Veterinária, Universidade de Évora, Núcleo da Mitra, Apartado 94, 7006-554 Évora, Portugal; [6]Faculdade de Medicina Veterinária, Universidade de Lisboa, Avenida da Universidade Técnica, 1300-477 Lisboa, Portugal; [7]David Marlin Consulting, P.O. Box 187, Cambridge, United Kingdom; monicademira@gmail.com

Received: 19 December 2019 / Accepted: 17 March 2020

RESEARCH ARTICLE

Abstract

The cardiac recovery index (CRI) is currently a key component of veterinary inspections to assess endurance horses metabolic status and fitness. Originally published by Ridgeway, it instructs veterinarians to subtract from the first heart rate (HR_1), collected when the horse is initially presented for examination, a second HR (HR_2), taken 1 min after the horse starts a 125 feet (38.1 m) out and back trot to assess gait. It is widely believed that an increase of more than 4 bpm from HR_1 might be an indicator of fatigue. The FEI rules instruct the veterinarians to start the stopwatch exactly 1 min after the HR_1 count instead of trot start, as described previously. The aims of this study were to investigate how time delays in the vet gate affect the HR_1 count and the CRI during endurance competitions, and to characterise and compare the time taken by veterinarians to measure the original version of the CRI ($tCRI_{RIDG}$) and the adapted CRI used in FEI endurance events ($tCRI_{FEI}$). Data from 972 veterinary inspections of horses that took place in different endurance competitions in three different venues were collected. There was no association between the time elapsed from entering the vet gate to the start of the HR_1 count or from the HR_1 count to the start of the trot-up and the HR_1 or the CRI ($P>0.05$). However, larger studies involving more venues and different layouts are needed to corroborate our findings and to characterise the sensitivity and specificity of the CRI regarding the baseline heart rate. Although this study did not show an influence of waiting times on the CRI, a reduced deviation from the mean observed across all veterinarians when using the original Ridgeway guidelines to calculate the CRI ($tCRI_{RIDG}$) seems to point a better time-wise consistency when this version is used.

Keywords: horse, equine, sports medicine, race, veterinary examination

1. Introduction

Competitive endurance rides take place over distances up to 160 km in one day against the clock. Similar to human marathons or trail races, they are designed to test the speed and stamina of a horse, in addition to the rider's capacity to conduct the horse across all kinds of terrain. However, equestrian endurance rides are not continuous races. The competition is divided into phases ranging from 16 to 40 km each, followed by a compulsory veterinary inspection which takes place in a designated area known as a 'vet gate'. A compulsorily rest period of between 20 and 60 min must take place thereafter. The assessment of the horses by the veterinary commission relies upon three criteria: heart rate recovery, metabolic stability and gait. The purpose of the assessment is to decide if a horse is fit enough to proceed to the next phase or, after the last phase (following completion of the race), if the horse has recovered within the prescribed time period (usually 30 min) to validate the competitor's placement (FEI, 2019).

Physiological recovery in endurance is commonly measured using the cardiac recovery index (CRI), which was first described by Ridgeway in 1991 and is therefore also known as the 'Ridgeway Trot'. Developed by a group of veterinarians involved with the American Endurance Conference (AERC) and based on unpublished observations, its aim was to assess the recovery of horses after passing the finish line by producing a recovery or CRI score. The CRI is currently recorded at each veterinary inspection (inspections after completion of each loop and re-inspections which take place during certain mandatory holds before the horse is allowed to start the next loop) to aid veterinarians in early identification of metabolic disorders and/or to support a decision to eliminate a horse.

The Ridgeway CRI method ($tCRI_{RIDG}$) as originally described (Ridgeway, 1991) consists of the subtraction of a second heart rate (HR_2) from the horse's original heart rate measurement when entering the vet gate (HR_1), taken exactly one minute after the horse starts a back and forth in-hand trot in a 125 feet alley or lane (38.1 m, adapted to 40 m by the FEI), where CRI = $HR_1 - HR_2$ (Mackay-Smith et al., 2016). A slightly modified version of the CRI method ($tCRI_{FEI}$) was published by Fielding *et al.* (2011) and embedded by the FEI with international endurance rules since then and at the time of publication (FEI, 2019)[1]. The $tCRI_{FEI}$ instructs veterinarians to start the stopwatch not when the horse starts the trot-up, but instead when they finish the HR_1 count (FEI, 2019).

Even if it seems intuitive that the end of HR_1 count matches the start of the trot-up, in reality this is often not the case. The flow of veterinary inspections might be disrupted by logistic factors inherent to endurance competitions, not only the layout of the vet gate, but also the number of veterinarians and trotting lanes available in relation to the number of competitors waiting to be examined, the space or absence of a delay between competitors determining the size of the pack to be examined, the overlap of competitions classes and the efficiency of judges and stewards directing the competitors. For instance, with the purpose of avoiding unfair and uneven waiting times between competitors from entering the vet gate to HR_1 count, more modern permanent venues and major championships will have special corridors for heart rate (HR_1) measurement at the entrance to the vet gate. Only then can horses with a heart rate meeting the 'pass' criteria (most commonly ≤64 bpm) proceed to a vetting lane. Thus, there is potential for a variable delay between HR_1 and HR_2, which could potentially affect the $tCRI_{FEI}$.

How the impact of time delays and the type of inspection may affect the CRI have to our knowledge not previously been described. Therefore, the aims of this study were to investigate: (1) how time delays in the vet gate affect HR_1 and the CRI measurements during endurance competitions; (2) to characterise and compare the time taken by veterinarians to measure the original $tCRI_{RIDG}$ and the adapted $tCRI_{FEI}$. Additionally, the study attempted to characterise the behaviour of the CRI for inspections after the loop (LoopInsp) and re-inspections within the hold (RInsp).

We hypothesised that logistics related to waiting times in the vet gate, namely time taken to HR_1 and from HR_1 to the start of the trot would affect the HR_1, HR_2 and therefore the CRI, and that the $tCRI_{RIDG}$ method would demonstrate more time-wise consistency in readings than the $tCRI_{FEI}$.

2. Material and methods

The study took place during three international endurance events between June 2018 and April 2019 in Loubejac (LOUB), Ligniéres (LIGN) and Uzès (UZES) in France. The study was approved by the FEI Veterinary Committee. Retrospective local climate information was obtained for each venue for the day of study from Weather Underground (www.wunderground.com). The nearest station with 30 min data reporting was identified. These were as follows: Loub – Bergerac Dordogne Perigord Airport, located 46 km WNW of Loubejac; Lign – Chateauroux-Centre Marcel Dassault Airport, located 36 km WNW of Lignieres; Uzès – Nimes Ales Camargue Cevennes Airport, located 25 km WNW of Uzès (Table 1). For data collection, an assistant was assigned to each officiating veterinarian and instructed to record times in a paper-based file of the following events during the veterinary inspections: (1) arrival of a horse at the veterinarian; (2) start and end of the count of HR_1; (3) start and end of the trot-up; (4) start of the clinical examination; (5) start and end of HR_2; and (6) end of the veterinary inspection. Data from the inspection after arrival from the loop and requested or mandatory re-inspections were used. Additionally, data from the timing systems and veterinary cards[2] were collected for each individual horse, including phase (loop), average speed, horse's recovery time (defined as the time lapse between arrival from the loop and entrance into the vet gate), time horse entered the vet gate, HR_1 and HR_2. For additional analysis HR_1 was grouped in four categories: HR_1 <56 bpm; 56≤ HR <60 bpm; 60≤ HR ≤64 bpm and >64 bpm. CRI was grouped in six categories: CRI <-8; -8≤ CRI <-4; -4≤ CRI <0; 0≤ CRI ≤4; 4< CRI ≤8; CRI >8. Since in France the CRI is not measured in the last vet gate after finish, the data from this stage were not used.

The CRI was calculated by subtracting HR_1 from HR_2 in accordance with the method described by Ridgeway

[1] The CRI was updated in the veterinary FEI rules in 2020 to the version originally described by Ridgeway as recommended by this paper.

[2] The clinical exam and the trot evaluation is registered at every veterinary examination in a specific card that accompanies each horse throughout the competitions.

Table 1. Mean (± standard deviation) and range (min-max) for shade temperature, relative humidity, wet bulb temperature, mean wind speed and rainfall, and overall reported condition for Loub, Lign and UzEs on days of data collection between 07:00 and 17:00.

	Shade temperature (°C)	Relative humidity (%)	Wet bulb temperature (°C)	Mean wind speed (km/h)	Rainfall (mm)	Overall condition
Loub	28±3 (22-32)	60±14 (43-88)	22±1 (20-23)	7±2 (2-9)	0±0 (0-0)	Fair
Lign	16±4 (8-21)	66±16 (46-100)	12±2 (7-15)	13±7 (2-24)	0±0 (0-0)	Fair
Uzes	15±3 (10-19)	41±6 (32-54)	9±2 (5-10)	23±2 (19-28)	0±0 (0-0)	Fair

(Ridgeway, 1991). The data were entered into a MS Excel® spreadsheet (Version 2016, Microsoft, Redmond, WA, USA) and the time elapsed from the horse entering the vet gate to reach an available veterinarian and to start the count of HR_1 (valid only for horses arriving from a loop) was calculated. In order to compare the classic $tCRI_{RIDG}$ definition of the CRI with the modified $tCRI_{FEI}$ version, the times elapsed between the start of the trot-up and the start of HR_2, and between the end of the HR_1 and the start of the HR_2 counts were calculated.

For further analysis, SPSS® version 22 (IBM Corp., Armonk, NY, USA) was used for descriptive and inferential statistics. Since data did not assume a normal distribution, a series of Kruskal Wallis analyses with post hoc Mann Whitney U tests identified if significant differences occurred between the variables recorded across events. A Wilcoxon signed-rank test assessed if calculated CRI values differed between the two tCRI methods. A Spearman correlation analysed if waiting times in the veterinary inspection impacted the HR_1 count and CRI values in the first inspections and re-inspections. Mann Whitney U analyses established if differences in CRI ($tCRI_{RIDG}$) and CRI ($tCRI_{FEI}$) existed between qualified and eliminated horses. Significance was set at $P<0.05$. HR and CRI variables are reported as mean ± standard deviation and median ± interquartile range (IQR) range.

3. Results

Data were obtained from 972 veterinary inspections (745 inspections and 227 re-inspections) of 352 horses. LIGN and UZES competitions took place in very similar weather conditions, but LOUB had higher temperatures (Table 1).

Data were collected from horses competing in open classes only (i.e. where there was no speed restriction), namely from 489, 53, 391 and 53 inspections from 80, 100, 120, and 160 km competitions, respectively. Overall, 22.5% (n=219) horse and rider combinations were eliminated, of which 18.7% (n=182) were for gait, 2.9% (n=28) were for metabolic reasons and 0.9% (n=9) for other reasons. More horses failed to qualify for the next stage in LoopInsp (n=179; 24%) than in RInsp (n=40; 18%). The HR_1 count was significantly lower in RInsp than in LoopInsp ($P<0.05$) and the CRI was significantly higher in RInsp than LoopInsp ($P<0.05$; Table 2). Significant differences ($P<0.05$) found between venues for HR_1 and CRI in LoopInsp are depicted in Table 3.

Mean, median, maximum and minimum time elapsed from horses entering the vet gate to start of the HR_1 count (for first horse inspections only VG to HR_1) and the time elapsed from the end of the HR_1 count to the start of the trot are reported in Table 4. No association could be found between the times elapsed from entering the vet gate to the start of the HR_1 count, from the HR_1 count to the start of the trot-up and of other stages of the inspection, and the HR_1 or the CRI ($P>0.05$). The time used to measure the $tCRI_{RIDG}$

Table 2. Descriptive statistics of heart rate (HR) and cardiac recovery index (CRI) in loop inspections and re-inspections.[1]

	HR (bpm)		CRI (bpm)	
	LoopInsp	Rinsp	LoopInsp	Rinsp
Mean ± sd	59±5	52±6	-1.0±4.0	0±4.3
Median ± IQR	60±6	52±8	-1.0±5.0	0.0±4.0
Minimum	38	34	-21	-20
Maximum	100	64	16	16

[1] sd = standard deviation; IQR = interquartile range; LoopInsp = loop inspections; Rinsp = re-inspections.

Table 3. Heart rate (HR) and cardiac recovery index (CRI) in loop inspections at the different venues.[1,2]

	HR (bpm)			CRI (bpm)		
	LOUB	LIGN	UZES	LOUB	LIGN	UZES
Mean ± sd	60.1±7.0	58.9±5.0	58.7±4.0	-0.5±5.0	-0.5±4.0	-2.0±4.0
Median ± IQR	60.0±6.0 [a]	60.0±5.0 [b]	59.0±6.0 [b]	-0.0±7.0 [a]	0.0±4.0 [a]	-2.0±4.0 [b]
Minimum	44	80	44	-20	-12	-21
Maximum	100	38	64	9	16	16

[1] sd = standard deviation; IQR = interquartile range.

[2] Different letters in a row show significant differences between venues.

Table 4. Time elapsed from entering the VG to HR_1, from HR_1 to the start of the trot and to perform CRI count according to $tCRI_{RIDG}$ and $tCRI_{FEI}$.[1]

	VG to HR_1	HR_1 to Trot	$tCRI_{Ridg}$	$tCRI_{FEI}$
Mean ± sd	1 m 28 s ± 1 m 18 s	6 s ± 11 s	1 m 0 s ± 13 s	1 m 5 s ± 16 s
Median ± IQR	1 m 10 s ± 54 s	3 s ± 7 s	58 s ± 12 s	1 m 3 s ± 14 s
Minimum	4 s	0 s	29 s	31 s
Maximum	20 m 4 s	3 s ± 5 s	2 m 30 s	4 m 33 s

[1] CRI = cardiac recovery index; sd = standard deviation; IQR = interquartile range; VG = vet gate; HR_1 = start heart rate 1 measurement; Trot = start of trot.

version versus the $tCRI_{FEi}$ version is represented in Table 4. Although significant differences were found between veterinarians (P<0.05) for both $tCRI_{RIDG}$ and $tCRI_{FEI}$, the $tCRI_{RIDG}$ method was more consistently conducted time-wise compared to the $tCRI_{FEI}$ method (Figure 1).

No association was found between HR_1 and CRI, except for the 160 km category where moderate and very strong negative correlations were found in the inspections after the loop (r^2 = -0.5; P<0.05) and in the re-inspections (r^2 = -0.9, P<0.01), respectively. However, there were a higher proportion of horses with CRI ≥4 in horses with HR_1 <60 bpm. For qualified horses, the distribution of the CRI was significantly different (P<0.05) among the heart rate categories of HR_1 <56 bpm, 56≤ HR_1 <60 bpm and 60≤ HR_1 ≤64 bpm in the first inspections after the loop and re-inspections, but not for eliminated horses (Figure 2). In the inspections after the loop horses within the categories of HR_1 <56 bpm (n=113; 20%), 56≤ HR_1 <60 bpm (n=234; 41%) had significantly higher CRIs than the category of 60≤ HR_1 ≤64 bpm (n=216; 39%). Within re-inspections the CRI was significantly higher for horses with a HR_1 <56 bpm (n=130; 70%), horses with 56≤ HR_1 <60 bpm (n=36; 19%), but not 60≤ HR_1 ≤64 bpm (n=17; 9%). The overall HR_1 for different CRI categories are presented in Figure 2.

No significant differences were found in the CRI of qualified versus eliminated horses in inspections or re-inspections (P>0.05). There was no association between recovery time, i.e. the time a horse takes to meet the heart rate criteria of

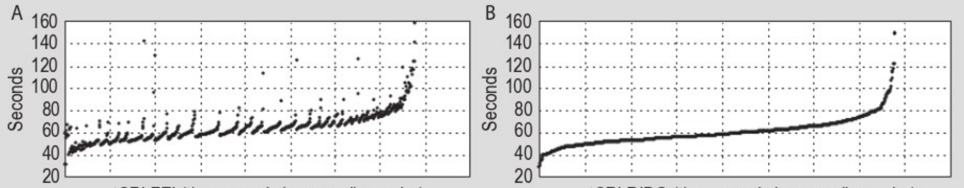


Figure 1. Time spent in cardiac recovery index (CRI) evaluation: original (Ridgeway) vs modified (FEI) definition.

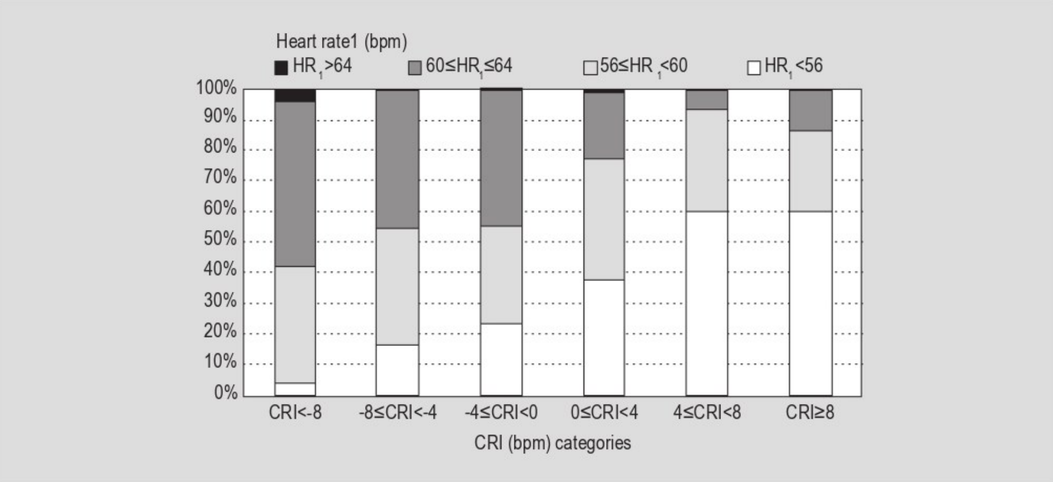


Figure 2. Distribution of heart rate frequencies according to cardiac recovery index (CRI) category in all horses in LoopInsp and RInsp.

less than or equal to 64 bpm before entering the vet gate, and the CRI (*P*>0.05). However, horses with CRI between -4 and -1 had shorter recovery times than horses with CRI between 0 and 4 or higher. HR_1 was not affected by recovery time, except in eliminated horses where recovery times were higher in horses with HR_1 >60 bpm.

4. Discussion

Interruptions to the flow of the veterinary inspections occur frequently (MM, personal observation) due to logistic reasons inherent to endurance competitions, which can result in involuntary and variable waiting times. In a busy vet gate, most commonly there will be a wait from the entrance into the vet gate to an available veterinarian in a trotting lane that is also usually assigned to perform the HR_1 count. The delay can also occur from the HR_1 count to the start of the trot-up in the following situations: when one or more veterinarian(s) are temporarily assigned to perform only the HR_1 count in a crowded vet gate; when the display of the gate is such that specific corridors at the entrance of the vet gate are assigned just for HR_1 count; occasionally when the veterinarian is called to participate in a voting panel (three veterinarians need to be anonymously consulted to eliminate a horse) in a different lane. Thus, delays can occur either from the entrance into the vet gate to the HR_1 count and/or from HR_1 count to the start of the trot-up. Less commonly, the trotting time will be delayed, because a horse might not be willing to trot straight away or not want to return; this can be due to factors such as young age, inexperience and/or tiredness.

To our knowledge, the impact of waiting times inside the vet gate in HR_1 and the CRI have not been investigated before. The findings of this study suggest that the CRI and the HR_1 are not significantly affected by variability in the waiting times between the phases of the veterinary inspection, outside of 160 km races. No association could be established overall, between venues or competition classes. This is important, because even if the time spent inside the vetting area does not influence the competition results directly, the CRI value, although not a fixed parameter for elimination *per se*, might influence the decision of the veterinary commission to eliminate a horse, call for a requested re-inspection or trigger a withdrawal by a competitor (Robert *et al.*, 2002).

As expected, the duration of the $tCRI_{RIDG}$ showed more consistency than the $tCRI_{FEI}$. The $tCRI_{RIDG}$ seems to be more appropriate, since it has fewer steps in between the counts, namely the waiting time between HR_1 and the start of the trot. Even if the CRI was not affected by time delays, our findings would advocate that the phrasing in the FEI rules should be changed to the $tCRI_{RIDG}$ since it better matches the reality of the vet gate logistics worldwide, especially at major championships.

Though used extensively worldwide, few peer-reviewed reports have evaluated the use of the CRI since the unpublished observations resumed by Slusher and Mackay-Smith, and Ridgeway in the proceedings of the Annual Convention of the American Association of Equine Practitioners (AAEP) in 1991 (Ridgeway, 1991; Slusher and Mackay-Smith, 1991). This work observed that fit horses quickly regained their resting heart rate compared to metabolically compromised horses, which would show more labile and higher heart rates after brief exercise (Slusher and Mackay-Smith, 1991). However, it is worth noting that the $tCRI_{RIDG}$ method was meant to be used 15 min after horses had passed the finish line (Ridgeway, 1991) in order to score the recovery category and not to eliminate horses during the phases of the competition. It was only later that it was adopted worldwide at recovery checks, renamed re-inspections by the FEI, then to become mandatory for every inspection, including the inspections performed immediately after each phase in races (FEI, 2019). Since then, veterinarians have been instructed to trot the horses immediately after HR_1 collection in all inspections instead of performing the metabolic examination first as previously prescribed. One of the peer-reviewed studies investigated the CRI in recovery checks (re-inspections) only before the last loop (Robert *et al.*, 2002) and the other report studied the risk factors of elimination which included the CRI, but did not specify which type of inspections were used for the study (Fielding *et al.*, 2011). The first study concluded that a CRI ≥4 had only a predictive value for elimination in the final vet gate if HR_1 >60 bpm. The second study concluded that the odds of elimination were increased with a CRI >4.

As expected, our findings were that HR_1 was higher in LoopInsp than RInsp. But they also showed that the CRI was higher in RInsp than LoopInsp. In-between venues, the CRIs were only different in LoopInsp, with LOUB, the venue with warmer weather, showing higher HR_1, but surprisingly not higher CRI. UZES had the highest CRI, which is most likely explained by the lower HR_1 counts recorded here.

Every endurance veterinarian and most riders/trainers worldwide are taught that a CRI higher than 4 bpm is a 'red flag' regarding the metabolic status of horses. However, the fact that the baseline HR_1 should be taken into consideration when interpreting CRI values is not as widely acknowledged. The 4 bpm value was based on the observations that sound horses were expected to return to (or be even lower than) the baseline measurement and that fatigued horses would have an elevation in the CRI of 4 to 20 bpm (Slusher and Mackay-Smith, 1991). Nonetheless, the behaviour of the CRI in relation to different HR_1 and the various timings after the conclusion of the phase was never statistically studied. Ridgeway (1991) highlighted that a rise of the CRI of 4 bpm when the HR_1 of a horse was close to resting values (40 to 48 bpm) had little significance and suggested that horses with higher HR_1 should be penalised more heavily. Both Fielding *et al.* (2011) and Robert *et al.* (2002) found a predictive value for elimination for a CRI >4, but Robert did not find a predictive value of the CRI *per se*, but only in association with HR_1 of more than 60bpm. Indeed, a recent non peer-reviewed publication stressed that the two criteria should be taken into account when interpreting the CRI: the magnitude of the heart rates, i.e. if both heart rates are greater than 60 bpm, and the difference between heart rate counts (Gillespie *et al.*, 2015). Our results also advocate that CRI >4 in horses with a HR_1 lower than 60 bpm lack accuracy and should be interpreted with caution. Independently of the outcome, horses with lower heart rates showed higher CRIs and horses with HR_1 between 60 and 64 bpm tended to have a lower CRI. Although a correlation between HR_1 and CRI could not be established, the proportion of horses with a CRI ≥4 was higher in horses with HR_1 <60 bpm (92%, n=55). Our findings suggest that the closer the HR_1 is to the resting heart rate of a non-exercised horse, the higher the CRI might be. For instance, a CRI of 8 in a horse that presents a HR_1 of ≤56 bpm might be a physiological rise and not be related to metabolic compromise. Therefore, best condition calculations or special classification protocols that take into account the CRI obtained in vet gates might not entirely reflect the metabolic status of a horse. Since we could not find any significant differences for the CRI between qualified and eliminated horses in inspections after the loop or re-inspections, and no associations between variables, a predictive value was not attempted. We could not find any correlation for any elimination, lameness or metabolic-related factors with CRI, in contrast to what has been previously reported (Trevillian *et al.*, 1997).

Endurance competitions have become very competitive in the last decade. Due to the awareness of the impact of the recovery time in the final results, but also in the market value of a horse, most top horses will enter the vet gate within 3 min. Our study did not show any association between recovery time and the CRI for qualified or eliminated horses. However, qualified horses registered lower recovery times and this most likely reflects the preparedness of those horses. We could not find those differences in eliminated horses. This supports the study by Younes *et al.* (2016), reporting that the recovery time is the most sensitive indicator of eliminations in endurance horses.

Elimination rates in endurance rides are high and mostly due to lameness (Bennet and Parkin, 2018; Marlin and Williams, 2018; Nagy *et al.*, 2014). The subjectivity of gait assessment often elicits confrontations between riders/trainers/owners and veterinarians (De Mira *et al.*, 2019). In a preliminary study using a portable system based on inertia sensors horses were instrumented after HR_1 count and before the start of the trot (Lopes *et al.*, 2018). Although not timed in this study, instrumentation of horses is reported to take 2 to 3 min (Equinosis, 2019), which means a delay between HR_1 and HR_2 counts. However, CRI performed according to the $tCRI_{RIDG}$ version would not be affected by the delay since the minute starts when the horse commences the trot-up.

We acknowledge that only three venues were studied and that this might not reflect the waiting times of busier competitions and/or different weather conditions. Therefore, larger studies involving more venues and different setups are needed to corroborate our findings. Also, the low numbers of eliminations were too small to determine the added value of the CRI to the different types of inspections and to determine significant cut-off values for specific heart rates. More studies are needed to characterise the sensitivity and specificity of the CRI regarding the baseline heart rate.

Although this study did not show an influence of waiting times on the CRI, reduced deviation from the mean was observed across inspecting veterinarians using the original Ridgeway guidelines to calculate the CRI, indicating more consistency. We would therefore suggest that the FEI veterinarian rules should be updated accordingly for improved coherence. Furthermore, the future use of instrumentation of horses with sensors for gait analysis in the future to determine/judge/evaluate lameness related eliminations should not interfere with the concept of the CRI.

Conflict of interest

The authors declare no conflict of interest.

References

Bennet, E.D. and Parkin, T.D.H., 2018. Fédération Equestre Internationale endurance events: risk factors for failure to qualify outcomes at the level of the horse, ride and rider (2010-2015). Veterinary Journal 236: 44-48. https://doi.org/10.1016/j.tvjl.2018.04.011

Equinosis, 2019. What can I expect during an objective lameness evaluation for my horse? Available at: https://equinosis.com/horse-owner-faqs-lameness-evaluation/.

Fédération Equestre Internationale (FEI), 2019. Veterinary regulations. Available at: https://inside.fei.org/fei/regulations/veterinary

Fielding, C.L., Meier, C.A., Balch, O.K. and Kass, P.H., 2011. Risk factors for the elimination of endurance horses from competition. Journal of the American Veterinary Medical Association 239: 493-498. https://doi.org/10.2460/javma.239.4.493

Gillespie, J.R., Kerr, J., Adamson, B. and Ellery, J., 2015. Adding to our understanding of cardiac recovery index at endurance rides. Endurance News: 19-24.

Lopes, M.A.F., Eleuterio, A. and Mira, M.C., 2018. Objective detection and quantification of irregular gait with a portable inertial sensor-based system in horses during an endurance race – a preliminary assessment. Journal of Equine Veterinary Science 70: 123-129. https://doi.org/10.1016/j.jevs.2018.08.008

Mackay-Smith, M., Bentham, B., Cohen, M., Nelson, T., Ridgway, K. and Steere, J., 2016. Guidelines for control judges and treatment veterinarians at AERC endurance competitions. Available at: https://aerc.org/static/upload/CJHandbook2016.pdf

Marlin, D. and Williams, J., 2018. Equine endurance race pacing strategy differs between finishers and non-finishers in 120 km single-day races. Comparative Exercise Physiology 14: 11-18. https://doi.org/10.3920/cep170027

De Mira, M.C., Santos, C., Lopes, M.A. and Marlin, D.J., 2019. Challenges encountered by Federation Equestre Internationale (FEI) veterinarians in gait evaluation during FEI endurance competitions: an international survey. Comparative Exercise Physiology 15: 371-378. https://doi.org/10.3920/cep180058

Nagy, A., Murray, J.K. and Dyson, S.J., 2014. Descriptive epidemiology and risk factors for eliminations from Federation Equestre Internationale endurance rides due to lameness and metabolic reasons (2008-2011). Equine Veterinary Journal 46: 38-44. https://doi.org/10.1111/evj.12069

Ridgeway, K.J., 1991. Inride veterinary examination, postride examination and judging of best condition. In: American Association of Equine Practioners Annual Convention (AAEP), Denver, CO, USA, pp. 815-826.

Robert, C., Benamou-Smith, A. and Leclerc, J.L., 2002. Use of the recovery check in long-distance endurance rides. Equine Veterinary Journal 34: 106-111. https://doi.org/10.1111/j.2042-3306.2002.tb05400.x

Slusher, S.H. and Mackay-Smith, M.P., 1991. Endurance ride veterinary control. In: Proceedings of the Annual Convention of the American Association of Equine Practitioners, Denver, CO, USA, pp. 793-805.

Trevillian, C., Holt, J. and Yovich, J.V., 1997. Evaluation of cardiac recovery index and clinicopathological parameters in endurance horses. Australian Equine Veterinarian 15(2): 83-88.

Younes, M., Barrey, E., Cottin, F. and Robert, C., 2016. Elimination in long-distance endurance rides: Insights from the analysis of 7,032 starts in 80 to 160 km competitions. Comparative Exercise Physiology 12: 157-167.

CHAPTER V: Vet gate trotting style improves subjective gait grading in endurance horses when compared to a lameness presentation style

Mónica Cardoso de Mira[1] | Constanza B. Gómez Álvarez[2,3*] | Rute Santos[1,4] | Orlando Fernandes[5] | Filipe Serra Bragança[6] | Massimo Puccetti[7] | Marie Rhodin[8*]

[*] CBGA and MR should be considered joint senior author.

Abstract

Background: To minimise the chances of failing to qualify due to lameness, it is anecdotally acknowledged that horses are trotted up differently at endurance competitions than during regular lameness workups.

Objectives: To investigate how a lameness workup (LWP) or a vet gate (VGP) trotting presentation style affected the agreement of experienced Fédération Equestre Internationale (FEI) endurance veterinarians using an 0-5 lameness grading scale (LGS) and the ABC endurance score systems when evaluating videoed trot-ups of slightly lame horses but likely to pass the vet gate evaluation.

Study design: Observational study with replicated measurements

Methods: Thirty-three horses slightly lame as defined by their owner/handler/vet, actively competing in endurance were trotted by experienced handlers in two presentation styles (LWP/VGP) and instrumented with a body inertial measurement unit system (IMUs). Six FEI 4-star veterinarians blindly evaluated videos of the horses. Speed was collected by a GPS watch. A generalised linear mixed model assessed the effect of presentation styles and speed, including trotting surface, handler on subjective scores. The predictability of 0-5 LGS and IMUs overall symmetry to determine fitness to compete was analysed using a Receiver-Operating Characteristic analysis. Subjective agreement was assessed by intra-class correlations.

Results: The interaction between VGP trotting style and speed elicited significant lower scores on the 0-5 LGS [F=2.49 (0, 5), p=0.03] and [F=2.62 (1, 4), p<0.01] and ABC endurance score [F=2.62 (1, 4), p<0.01]. Moderate and higher speeds generated lower scores in the VGP (p<0.01) than the LWP, but there was not a significant difference between trotting styles in slow speed. LGS recorded excellent predictability for gait-related failing to qualify (AUC: 0.93, CIs: 0.90 – 0.95), with a lameness grade > 1 on the 0-5 LGS (88% sensitivity and 90% specificity) proposed as a threshold for failing a horse. 25% of the trot-ups deemed fit to compete showed severe lameness (>18mm front or 9mm hindlimb) with the IMUs. Veterinarians scored lower grades more frequently in hind than in forelimb lameness (p=0.012), while there was no difference in the IMUs analysis.

Main limitations: Video-analysis does not entirely reflect live evaluations and is known to reduce ICC among raters. Handlers' order was not randomized.

Conclusions: VGP style coupled with higher trotting speeds improves lameness scores when evaluated by veterinarians using both the 0-5 LGS and ABC endurance scales**.** The 0-5 LGS showed excellent predictability for gait-related failure to qualify; however, many horses deemed fit to compete

were objectively severely lame. Front vs. hind limb lameness are unequally scored by experienced FEI veterinarians.

Introduction:

Compulsory veterinary inspections to determine if horses are fit to continue to compete take place before and after every phase of endurance competitions and after the last phase, to ensure horses are sound enough to earn the classification obtained when crossing the finish line. Inspections are performed by a veterinary commission (VC) in an assigned area called the vet gate (VG) (FEI, 2021a; b). A decision to fail a horse must be legitimated by a panel composed of three veterinarians voting anonymously, called upon the request of one veterinarian. Horses are eliminated either for metabolic reasons or lameness, and the cause is published with the event results. The intensity and duration of the effort, along with the number of veterinary controls to prevent severe injuries result in high rate of horses failing to qualify in endurance when compared to other equestrian disciplines. Lameness is the primary reason for a horse failing to qualify to continue competing in endurance, accounting for approximately 70% of all eliminations (Marlin and Williams, 2018).

Despite the importance of gait assessment, guidelines offered to official veterinarians by the FEI rules on how to interpret and establish a threshold to fail a horse for lameness in endurance competitions are scarce. The only definition provided in the veterinary rules (FEI, 2021b) is that: 'any pain-induced irregular gait in a 40 m forth and back straight-line trot-up observed without a limb flexion or deep palpation should lead to a failure to qualify'. No further details are given, but veterinarians are required to classify gait as fit to compete, grading as either A, i.e., sound or B, i.e. a not consistent irregular gait; or not fit to compete, grading as C, i.e. clearly consistently lame. Apart from additional indications received by email once in 2013 (FEI, 2013) with the same brief presentation at mandatory FEI courses and, more recently, in a non-open access published study funded by the FEI (Bennet *et al.*, 2020), official veterinarians have no access to other written guidance, relying mostly on their individual knowledge, and therefore make decisions based on their own or shared experiences with other OV's.

Within national federation endurance rules for countries with a longstanding tradition in endurance, for example the United States, Australia or South Africa, scores are explained more in detail (AERA, 2020b; AERC, 2016; ERASA, 2018), and further advice is given on the importance of being consistent in classifying gait throughout competitions (AERA, 2020b) (Information resumed at Supplementary Table S1). The American Association of Equine Practitioners (AAEP) 0-5 lameness scale, used by many practitioners in their daily practice worldwide, is used to score gait in American Endurance Riding Conference (AERC) endurance rides, whose rules state that grades 1 and 2 translate to a horse being fit to compete. In a recently published paper, only a 58% agreement was found in judging sport horses fit or not fit to compete by experienced FEI veterinarians, but no relationship with an ordinal score was investigated (Serra Bragança *et al.*, 2020). Inter- and intra-rater agreements for equine lameness assessment are reported to be low, especially in mild lameness, e.g. ≤1.5 (AAEP)(Keegan *et al.*, 2010). However, the ABC scale used in other endurance rules (AERA, 2016;

ERASA, 2018; FEI, 2021a) incorporates more parameters than gait regularity, such as rhythm, elasticity and vigour, and more subjective and colloquial characterisations such as “a horse not showing itself off very well” (Bennet *et al.*, 2020). To fail due to lameness (C score), a horse has to be clearly lame (Bennet *et al.*, 2020). To the author’s knowledge, no previous studies have addressed if a C score is interpreted by veterinarians as a "2" or a "3" grade, when using a 0 to 5 lameness grading scale. Moreover, it is tacitly acknowledged by the endurance community that a routine lameness workup trot-up presentation (LWP) differs from a presentation performed at an endurance competition (i.e., vet gate) (VGP), where the handler aims to minimise the chances of the horse showing any lameness. Furthermore, often a specific handler with better trotting skills is assigned to present a horse, especially when the voting panel is called. As expected, the handler trotting a horse was considered by 94% of FEI endurance veterinarians as the most critical factor disturbing lameness evaluation in competitions (de Mira M.C. *et al.*, 2019). Handlers might intentionally or unintentionally rise the trotting speed, which has been reported to improve the lameness score assigned by evaluators (Starke *et al.*, 2013), or interfere with the horse’s head position by pulling, lifting or other adjustment to avoid a consistent head nod(de Mira M.C. *et al.*, 2019).

The aim of this study was to investigate the repeatability and agreement of six experienced FEI endurance veterinarians’ scoring using a 0-5 lameness grading scale (LGS) and the ABC endurance score system, when blindly evaluating videoed trot-ups presented in a lameness workup (LWP) or a vet gate (VGP) presentation style. It was hypothesised that a VGP would be performed at a higher speed and generate lower lameness scores than LWP. Additionally, the study aimed to compare the 0-5 LGS and the collected objective data, interpreted by previously reported guidelines, with the ABC endurance score. A threshold in the 0-5 LGS and objective overall symmetry calculations for the outcomes PASS (A and B scores) and FTQ (C score) was also investigated.

Material and methods:

Horses

Thirty-three horses (11 mares and 22 geldings) actively competing in endurance, deemed by the owner or stable's veterinarian to have at some point shown a slight lameness, but likely to pass a vet gate evaluation (B trot) from five stables in Portugal (stables 1, 2 and 3) and Dubai (stables 4 and 5) were included. Exclusion criteria were lameness of 3 or more out of 5 observed by the first two authors, 0-5 lameness grading scale (0-5 LGS), with a "0 grade" corresponding to no lameness and a 5 grade to a no weight-bearing lameness, which excluded one horse. Information regarding breed and age were retrieved from the FEI data base (https://data.fei.org/). Eighteen horses were registered as purebred Arabians, six as Part Bred Arabians, two as Anglo-Arabians, and six as unknown breed. The mean (±SD) and median (±IQR) age of horses was 10.8(±3.2) and 10.0(±3.0) years, respectively.

Handlers and trot-up instructions

Experienced handlers accustomed to trot at endurance competitions or working at professional endurance stables were assigned and asked to trot the horses back and forth for visual lameness examination in two different styles: as a regular lameness workup presentation (LWP) for the stable's veterinarian and as a vet gate presentation (VGP) for an official endurance veterinarian. No further instructions were given since, as expected, all handlers elected to trot horses differently in the two scenarios. Each horse was trotted by three different handlers using the two trotting styles twice and a total of 13 handlers were used in the study. Distribution of handlers per stable and horse can be consulted in Figure 1.

Trot-ups and inertial sensor-based gait analysis system

In the three first stables (Portugal), horses were trotted in a firm dirt terrain, in contrast to the last two stables (Dubai) that had a specific built-in hard and even trotting surface lane. The turning point was set at 40 m from the starting point. After a dynamic visual examination of a back-and-forth trot-up by the first author, the horses were instrumented with four inertial measurement units (IMUs)[a] (Lameness Locator): attached to the poll with a felt head bumper, taped to the midline of the withers, taped between the sacral tuberosities, and a fourth unit (consisting of a uniplanar gyroscope to measure the rotation of the limb in the sagittal plane) taped to the dorsum of the pastern of the right frontlimb. After habituation to the trot-up area, each horse was trotted twice in hand, back and forth in a straight line in each trotting style, by each of the three handlers, making a total of 12 trials per horse (Figure 1).

Handler order was randomised, but the trotting style order was maintained. Each trial was registered by the IMU system and simultaneously recorded by one video camera mounted on an automated zooming robot[b]. Data from the IMUs were wirelessly transmitted to a computer tablet and saved for analysis. A GPS Polar® watch was placed in the wrist of the handler to record the trotting speed of the horse and handler. Data were collected at a frequency of 1Hz and retrieved into the Polar Flow® software, which allowed the identification of the distance covered in the trot-up and speed. Data were computed into Python 3.7.4, for uniformisation, all trot-ups were trimmed to a (forth) starting and (back) ending speed of 6.9 km/h, the reported speed at which Arabian horses start to trot (Griffin *et al.*, 2004).

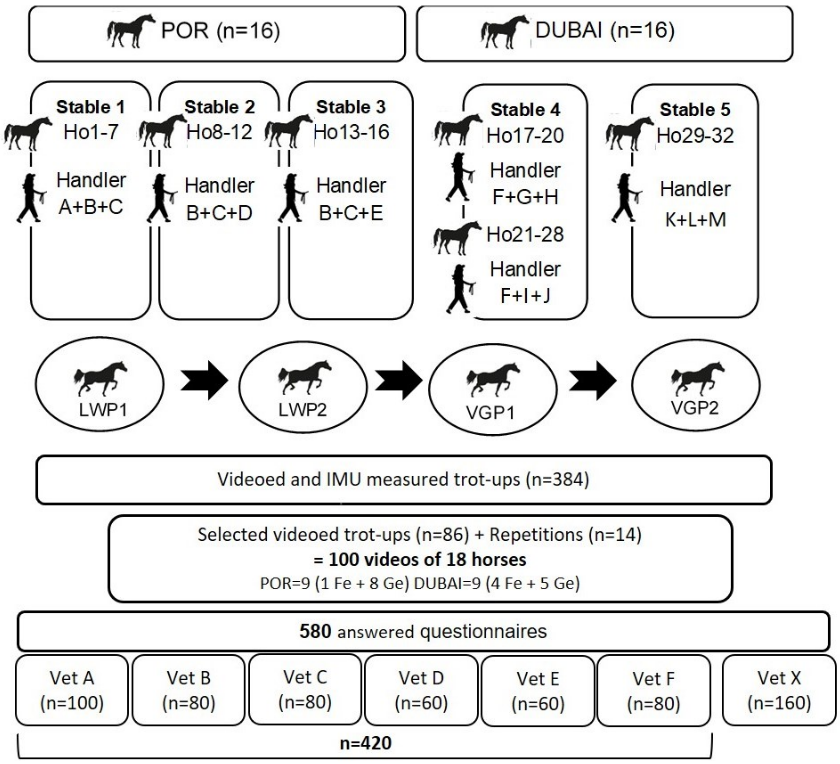


Fig. 1: Study Design. Distribution of handlers and horses by stable. After habituation to the trot-up area, each horse (n=32) was trotted twice in hand, back and forth in a straight line in each trotting style, by each of the three handlers, randomly assigned, making a total of 12 trials per horse. A total of 384 videoed and objectively measured trot-ups were obtained. Of those, 86 videoed trot-ups with objective measurements were selected to enter the study and of those 14 were selected to be repeated, making up a final sample of 100 videos, each sent with a questionnaire to six FEI 4* veterinarians (Vet). A total of 520 answered questionnaires (scores) were generated, where 420 identified the Vet, and could therefore be used for Intra-Class Correlation (ICC) analysis, but the remaining 160 did not. Lameness Work-Up Presentation. VGP: Vet Gate Presentation. n (number or trot-ups or scores).

Speed collection and analysis

Video-analysis

Each trial was simultaneously recorded with one camera mimicking the standing position of a veterinarian in a vet gate, e.g., trotting away and back, for further subjective lameness evaluation. As each horse was trotted by three different handlers, and each handler trotted each horse twice at two different styles for security of sampling, i.e., each horse was trotted 12 times; a sample of 384 videos was obtained. Any video with low quality and or with missing objective measurements data were excluded. The first author then randomly selected an equal amount of sound, front and hindlimb lameness videos. Furthermore, the best quality video including trot-up repetitions (same handler and trotting style) was selected, assuring that either the handler or the trotting style were different. This resulted in a sample of 86 trot-ups from 18 horses, of which 2 presumably were sound, 7 and 5, front and hind limb lame, respectively, and the remaining multiple limb lame, e.g., when a one-limb lameness was not obvious. Of those, 14 videos were randomly selected to be blindly repeated for intra-rater agreement analysis.

The selected trot-ups for video-analysis were then assigned randomly using online software (www.random.org) to 5 sets of 20 trot-ups, including the 14 repeated trot-ups. Each set was then embedded into a Google Forms questionnaire (Supplementary Material S2) and sent to six FEI 4-star veterinarians worldwide, two from Europe, one from Australia, one from South Africa, one from South America and one from North America. The veterinarians were asked to judge the trot-ups as if in an endurance competition, setting grading gait into A, B and C grades, and using a 0-5 lameness grading scale (0-5 LGS). They were also asked whether they would call for a voting panel for that horse. They were not informed of the two trot-up styles performed by the handler. For each trot-up evaluation, means±S.D., medians±IQR and ranges were calculated for both grading systems with the ABC endurance score being coded into a numerical score (A=0, B=1, C=2). The assessments medians were further grouped into PASS (fit to compete A/B scores) and FTQ (Fail to Qualify: C scores) outcomes.

Inter- and intra-observer scoring agreement

Inter- and intra-observer agreement of 420 observations (scores) attributed by six raters were calculated using intraclass correlations (ICC) estimates and their 95% confident intervals. A custom-made code in MATLAB version 7 (The MathWorks Inc. Natick. Massachusetts), based on a mean rating using multiple raters, absolute-agreement, 2-way random-effects model was used. The level of reliability was interpreted according to Koo and Li (2016), where values less than 0.5 are indicative of poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.9 indicate good reliability, and values greater than 0.90 indicate excellent reliability.

Objective analysis

Mean and standard deviation of the vertical movement asymmetry of the head (HDmin and HDmax), withers (WDmin and WDmax) and pelvis (PDmin and PDmax) were calculated per trial as previously described (Persson-Sjodin *et al.*, 2019). Each trial was classified into single limb lameness or multi-limb lameness, including patterns definitions as impact or push-off, following the guidelines reported by Reed et al. (Reed *et al.*, 2020) (Supplementary Material Table S2). Frontlimb and hindlimb movement asymmetry was considered to be significant when the mean head vector sum (VS) (square root of [HDmax2 + HDmin2]) was > 8.5 mm with a standard deviation (SD) of < 120% and absolute mean pelvic motion asymmetry (PDmin and/or PDmax) >3mm with a SD < 100% of the mean values (Reed *et al.*, 2020). To distinguish primary frontlimb lame horses from horses with compensatory head movement asymmetries in horses with ipsilateral movement asymmetries, the sign of the withers Min Diff as previously reported by Rhodin et al. (Rhodin *et al.*, 2018) was used. When signed similarly to HDmin, the principal lameness was attributed to that frontlimb, and when signed differently to the ipsilateral hindlimb (Pfau *et al.*, 2018; Rhodin *et al.*, 2018). HDmin and PDmin were used to categorise the severity of frontlimb (sound <6 mm, mild 6–12 mm, moderate 12–18 mm, severe >18 mm) and hind limb (sound <3 mm, mild 3–6 mm, moderate 6–9 mm, severe >9 mm) gait asymmetry, as earlier described by Hammarberg et al. (Hammarberg *et al.*, 2016b). Additionally, the overall symmetry (OS) was calculated as VS of the head/2 + VS of PDmin and PDmax, as reported by Serra Bragança et al. (Serra Bragança *et al.*, 2020).

Statistical data analysis

SPSS v.23.0.0.3 (IBM, Armonk, New York, USA) was used for descriptive and inferential statistics. Interpretation of objective data for soundness, lameness severity and lame limb(s) were aligned with the medians of both scores for comparison. Trot-ups graded as C were classified as FTQ, as opposed to A and B medians that were classified as fit to continue or PASS. Frequencies were counted and cross-tabbed with the 0-5 LGS. The distribution of 0-5 LGS and ABC endurance scores between the front and hind limbs were analysed by a Mann-Witney U test.

A generalised mixed model using a Poisson distribution with log link function was used to estimate the impact of presentation styles, speed, handler and surface on the subjective scores (0-5 LGS and ABC endurance score), due to the non-normality of data as determined by the Kolmogorov-Smirnoff test. Only 420 of the 580 answered questionnaires to the 100 video assessments whose raters' identity was known entered the analysis. Considering that horses in Dubai were trotted on built-in hard trotting lanes and in Portugal trotted on natural dirt lanes, country was coded as trotting surface and used to analyse any effect of trotting on results. A positive Spearman's rank correlation coefficient identified speed as having a moderator effect over the LWP (rho=0.3, p=0.01) but not the VGP style and therefore the interaction between presentation style and speed was included as an additional fixed factor. Moderating quantification was assessed using the Process macro version 3.5 (https://www.processmacro.org/). A regression analysis excluded multicollinearity between the interaction and the main factors. Horse intercept and veterinarian entered the model as the main random effects identified by an ICC analysis. The fitness of the model was assessed by a backward variables' elimination, and ultimately trotting surface was removed for non-significance and to yield a lower Akaike Information Criterion (AIC).

ROC analysis

The predictability of the classification systems 0-5 LGS and OS was assessed using Receiver-Operator Characteristic (ROC) analyses for the outcomes PASS and FTQ. Subsequent evaluation of sensitivity and specificity coordinates of each curve identified potential threshold values for FTQ for each system. Significance was set at $p < 0.05$.

Sample size

The study used a convenience sample that included the available horses at each visited stable. Subsequently, a power calculation free software (G*Power 3.1.9.2. Heinrich Heine Universität) showed that a total sample size of 176 observations was required for a difference of 1 grade in the 0-5 LGS and the ABC endurance scores between the trotting styles with a standard deviation of 0.5, based on a power of 0.95 and alpha of 0.05. Our sample size consisted of 580 observations (scores) used for descriptive statistics and ROC analysis, 420 observations used for ICC and 210 observations used for the GLMM and, therefore, considered appropriate.

Results

Subjective gait assessment

Of the 33 horses, one mare was excluded in the trot-up for a 3 out of 5 lameness. After a quality screening, 100 videos from only 18 horses were selected, being the remaining 14 horses excluded from this study. Six 4-star FEI endurance official veterinarians completed 580 questionnaires, with a range of 40 to 100 evaluated trot-ups (VetA:100, VetB:80, VetC:80, VetD:60, E: 60 and VetF: 40), as not all veterinarians responded to all sets of videos. Of those, 300 and 280 corresponded to a LWP and a VGP trot-up styles, respectively. There were 160 completed questionnaires with no identification of the respondents due to an error in the initial forms, and therefore 420 videos were used for the intra- and inter-rater agreements of veterinarians. The proportion of grades of the 0-5 LGS versus qualification of the trot-ups can be seen in Figure 2 and 3, respectively. Using the observers' median values for each evaluation, of the 86 trot-ups, 38, 30, 13 and 5 trot-ups were graded 0, 1, 2 and 3 out of 5, and 32, 36 and 18 were graded A, B and C, respectively (Supplementary Table S3).

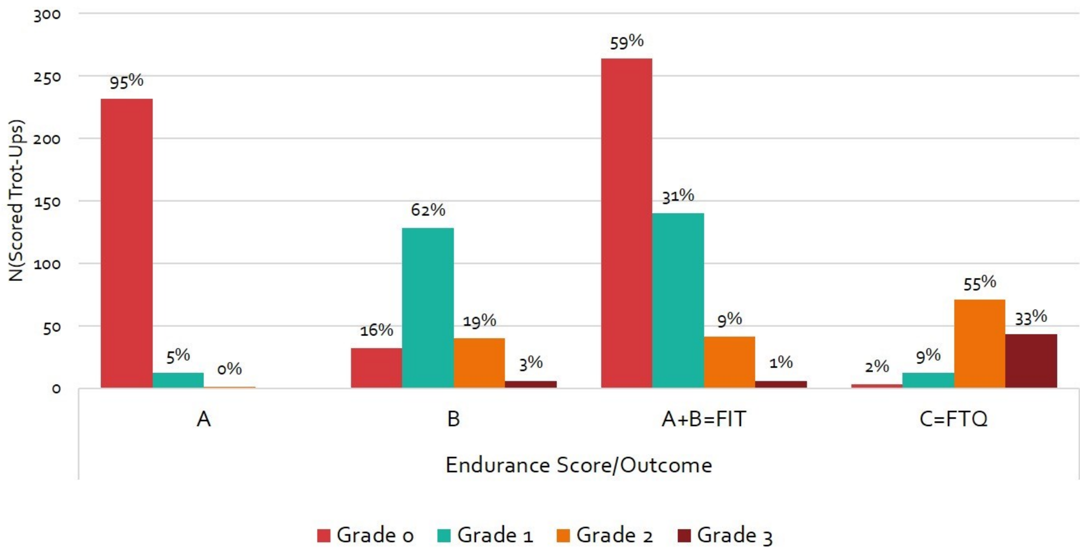


Fig. 2: Distribution of 0-5 lameness scale grade (LGS), plotted against veterinarian's trot-up scores (n=580) deemed to PASS (fit to compete or A+B scores) and fail to qualify (FTQ or C scores). % (relative percentages for grades A, B, C, PASS and FTQ)

In the ROC analysis, the 0-5 LGS recorded excellent predictability for FTQ (AUC: 0.93. CIs: 0.90 – 0.95), with a lameness grade > 1 out of 5 representing the threshold for FTQ at a sensitivity of 88% and specificity of 90%. Applying this threshold and the endurance subjective fit to compete decision as the gold standard, five horses would have recorded five false negatives (FTQ under the cut-off, therefore passing when they should have failed), and 24 horses would be recorded as false positives (PASS above the cut-off, passing when they should have failed). See Figure 4.

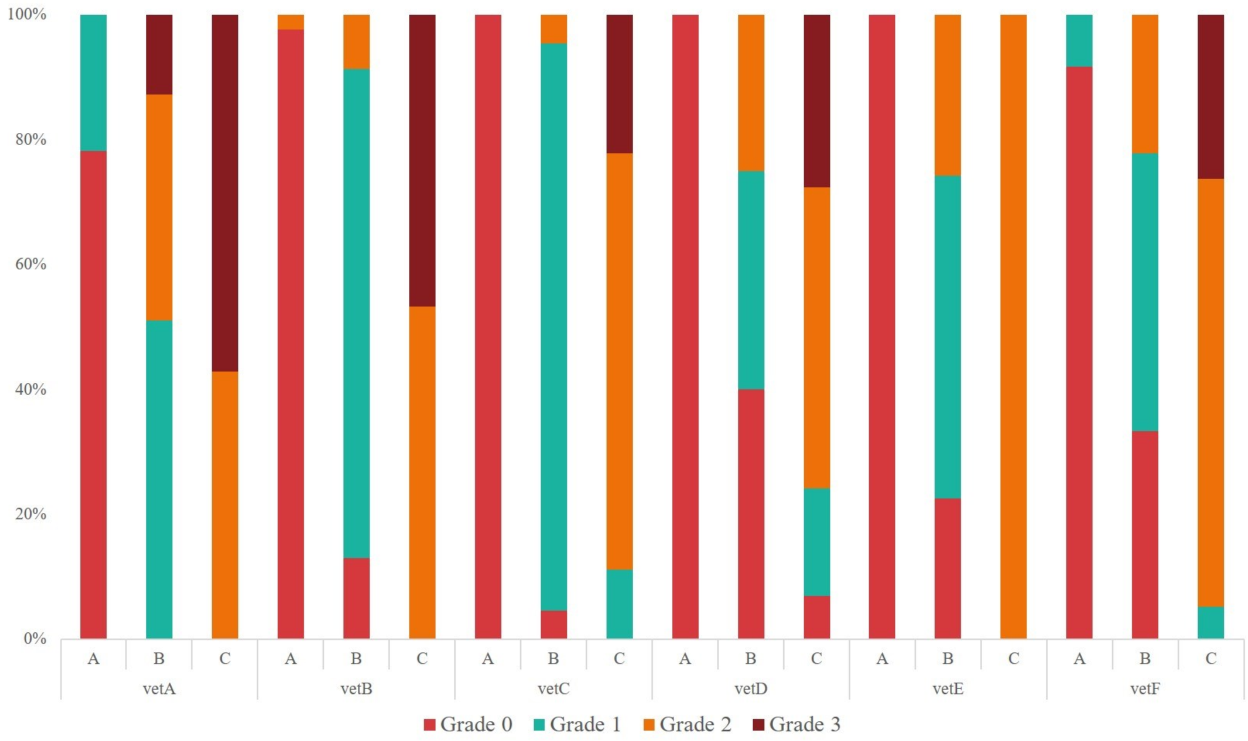


Fig. 3: Distribution (%) of 0-5 lameness scale grades given to trot-ups scored A, B and C after 580 evaluations (scores) by video-analysis per veterinarian (Vet A-F).

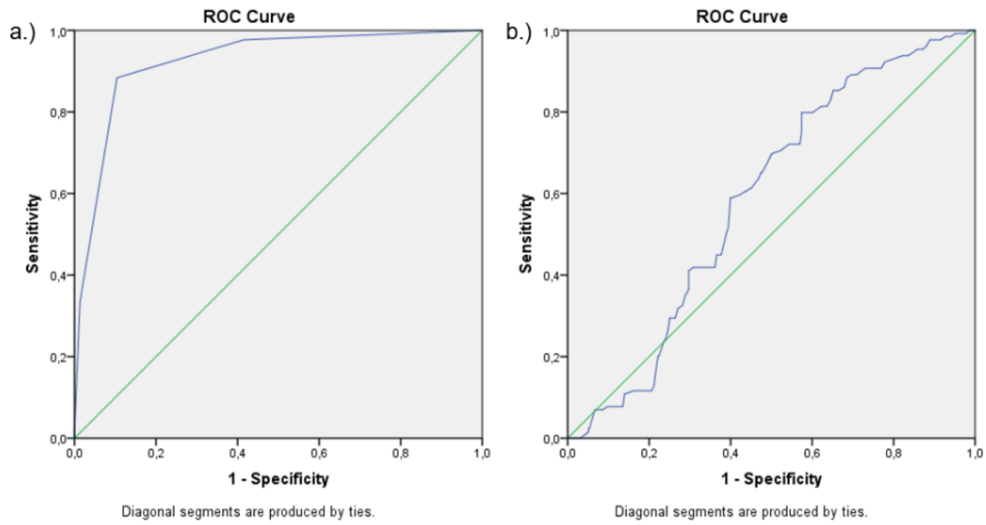


Fig. 4: Random Operating Characteristic (ROC) analysis of the 0-5 lameness grading scale (LGS) scores and (b) Overall Symmetry (n=580), plotted against the outcome PASS (fit to compete or A+B scores) and FTQ (Fail to Qualify or C scores).

Objective analysis

The objective measurements did not detect movement asymmetry in 26 out of 86 trot-ups. Only one horse consistently reported as symmetric across all six trot-up objective evaluations (Table S4). The remaining 63 trot-ups recorded asymmetry in 43 and 17 in one and two limbs, respectively. The wither's WDmin was concordant (same sign) to the asymmetric front limb in 20 of the 22 front limb asymmetries. Of the hind limb asymmetries there were 10 wither's WDmin discordant (opposite sign) and 11 concordant to the hindlimb asymmetry. The most common identified pattern in multiple limb asymmetries was a primary frontlimb impact asymmetry with a compensatory contralateral impact and push-off hindlimb lameness(n=7) or a contralateral only push-off hindlimb asymmetry (n=3), according to the second law of sides. Of those, trot-ups in the first pattern all showed a withers' WDmin and WDmax concordant direction (same sign), but in the second pattern, one out three trot-ups showed a discordant wither's WDmin direction and a concordant WDmax. The second most identified pattern was a primary push-off hindlimb with a compensatory ipsilateral impact frontlimb asymmetry (n=3) in accordance to the first law of sides. Even if all three trot-ups showed a concordant MinDiff and MaxDiff withers sign, they were classified as a primary hindlimb lameness. A front limb asymmetry with an ipsilateral impact asymmetry was identified in two horses with a concordant MinDiff and MaxDiff withers sign to the front asymmetry and therefore the primary asymmetry was attributed to the front limb. Distribution of trot-up's objective results among horses can be consulted in Supplementary Table S4.

When confronted to the subjective trot-ups evaluations, there were 68 trot-ups deemed to PASS (A and B scores) and 18 deemed to FTQ (C score) (Figure 5). One sound and one mild asymmetry trot-up out of 26 out of 11, respectively had a FTQ outcome, whereas 10 mild, 18 moderate and 15 severe asymmetries, out of 11, 23 and 26 trot-ups subjectively had a PASS subjective outcome.

The ROC analysis of OS showed a poor predictability for FTQ (ROC: 0.69, CIs: 0.58 – 0.80) with the best threshold found at 15.10 mm with a 72% sensitivity for a 65% specificity (Figure 4).

Front versus hindlimb lameness' classification

Of the 580 answered questionnaires where a limb was identified as the cause of lameness (n=249), veterinarians identified more often a front limb lameness (n=138; 55%) than a hindlimb lameness (n=74; 30%) or multiple lameness (n=37; 15%). There was not a significant difference in the 0-5 LGS or ABC score between front and hindlimb (Figure 5). Of the 86 objectively assessed trot-ups there were 38, 36 and 26% were attributed to a front, hindlimb or multiple limbs, respectively.

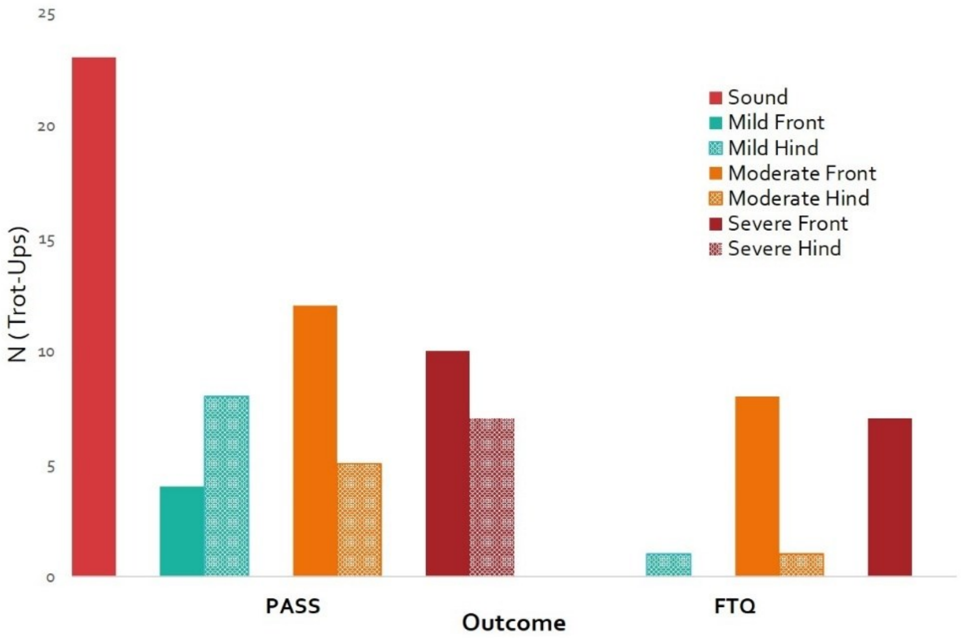


Fig. 5: Distribution (%) of trot-ups' (n=86) subjective evaluation calculated from the median of veterinarians' scores in PASS (fit to compete or A+B scores) and fail to qualify (FTQ or C scores), plotted against the objective evaluation in sound, mild, moderate and severe, front and hind gait asymmetry. Absolute numbers of evaluated trot-ups are represented inside the columns.

When reclassified according to the law of sides, there were 60 and 40% front and hindlimb asymmetries. The severity degree distribution was not significantly different between front and hindlimb in both subjective scores, but the objective assessment identified the frontlimb asymmetries as significantly ($p<0.00$) more severe than the hindlimb asymmetries (Figure 6).

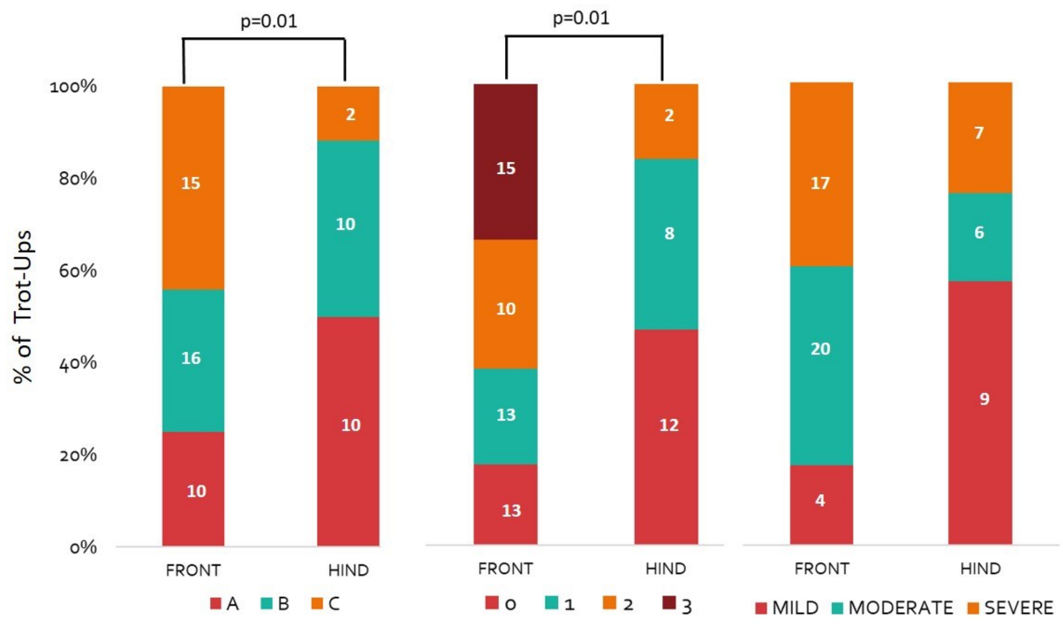


Fig. 6: Distribution of trot-ups' (n=86) subjective evaluation calculated from the median of veterinarians' scores using f the ABC endurance score and 0-5 lameness grading scale, and objective evaluation lameness classification (mild, moderate and severe) in front and hind limbs defined as asymmetric (n=60) by the objective measurements. P-values indicate a significant difference between front and limb lameness evaluations. % (relative percentages for PASS and FTQ)

Agreement between raters and repeatability of scores

Total agreement (100%) between raters for lameness grading was only found on six out of 86 occasions: four times for Grade 1 and score A, once for score C and only once in identifying the lame limb (front lameness). Intraclass Correlation Coefficients (ICC) were moderate for lameness grading scale 0-5, ABC endurance score, panel calling and limb identification. When compared with the LWP (0.53, p=0.05), the VGP style (0.70, p=0.00), elicited a higher 0-5 LGS agreement. However, the ABC endurance score agreement was higher for the LWP (0.81, p=0.00) and lower for the VGP (0.61, p=0.02). The VGP style recorded a better agreement for the need to call a panel (0.79, p=0.00) and limb identification (0.61, p=0.02) when compared to the LWP (0.71, p=0.00 and 0.48, p=0.01, respectively) (Table 1). ICC intra-rater values for blind repeated video-analysis were poor for lameness grading using the 0-5 LGS scale [ICC: 0.34, 95% CI (-0.22, 0.73)] and moderate for the ABC endurance scale [ICC: 0.55, 95% CI (-0.03, 0.84)].

Table 1: Inter-observer agreement using intra-class correlation (ICC) single measures of subjective assessment of lameness grade scale 0-5, ABC endurance score, whether a voting panel should be called and lame limb identification of all trot-ups, and separately, for lameness work-up presentation (LWP) and vet gate presentation (VGP) styles.

			95% Confidence Interval					
ICC -All trot-ups	Classification	Intraclass Correlation	Lower Bound	Upper Bound	F	*df*1	*df*2	p-value
Grades 0-5	Moderate	0.71	0.56	0.82	2.87	58	80.56	0.00
ABC Score	Moderate	0.60	0.38	0.74	2.01	58	80.56	0.00
Panel call	Moderate	0.68	0.51	0.79	2.48	58	112.28	0.00
Limb	Moderate	0.66	0.50	0.78	2.33	58	157.31	0.00
ICC - LWP		Intraclass Correlation	Lower Bound	Upper Bound	F	df1	df2	P
Grades 0-5	Moderate	0.53	0.20	0.76	1.84	25	41.81	0.05
ABC Score	Good	0.81	0.66	0.91	4.38	25	50.89	0.00
Panel call	Moderate	0.71	0.48	0.86	2.94	25	50.89	0.00
Limb	Poor	0.48	0.10	0.73	1.47	25	92.89	>0.09
ICC - VGP		Intraclass Correlation	Lower Bound	Upper Bound	F	df1	df2	P
Grades 0-5	Moderate	0.70	0.40	0.86	3.02	21	43.21	0.00
ABC Score	Moderate	0.61	0.28	0.82	2.14	21	43.21	0.02
Panel call	Good	0.79	0.62	0.90	3.94	21	64.08	0.00
Limb	Moderate	0.61	0.29	0.82	1.95	21	87.90	0.02

F: F-Test. *df:* degrees of freedom

Effect of trotting speed on subjective evaluation

The average speed (Table 2) with the cut-off set at 6.9km/h (or 2.7m/s) was significantly higher ($p=0.001$) for the VGP when compared to LWP style.

Effect of trotting styles on subjective assessment scores

A significant moderating effect of speed on trotting styles impact in the 0-5 LGS [t=3,02 (-0.51, 0.11), $p < .003$] and ABC endurance scores [t=3,02 (-0.51, 0.11), $p < .003$] was demonstrated and, therefore, the interaction trotting style*speed was introduced in the final GLMM model. The standardised slopes of interchanging moderating effects between the two interaction variables is depicted in Figure 7. At lower speed level an effect between trotting styles was not found. However, at higher speeds there was a significant difference ($p<0.01$) between the trotting styles in both subjective scores.

Table 2: Difference between maximum speed and average speed of lameness workup (LWP) and vet gate presentation (VGP) styles expressed in km/h and m/s. IQR: Interquartile Range. § Estimated marginal means. SE: Standard Error.

	Median±IQR (Range)		Mean§±SE		P-Value	95% Confidence Interval			
	LWP	**VGP**	**LWP**	**VGP**		LWP Lower	LWP Upper	VGP Lower	VGP Upper
Speed									
(km/h)	10.2±2.90 (8.5-12.1)	11.4±1.45 (9.0-13.7)	10.5±0.23	11.4±0.23	< 0.00	9.94	11.0	10.87	11.92
(m/s)	2.8±0.215 (2.4-3.4)	3.17±0.40 (2.5-8.81)	2.9±0.06	3.0±0.06		2.76	3.06	3.02	3.31
Max speed									
(km/h)	12.4±1.60 (9.6-22.3)	14.4±2.18 (10.4-19.1)	13.0±0.41	14.8±0.41	0.001	12.13	13.91	13.89	15.66
(m/s)	3.4±0.44 (2.7-6.2)	4.00±0.60 (2.9-5.3)	3.60±0.1	4.1±0.11		3.37	3.86	3.86	4.35

However, there were different effect of the analysed factors on the subjective scores. When considering the LWP, the VGP style factor on its own generated higher subjective scores, e.g. [0-5 LGS: F=2.49 (0, 5), p=0.03] and [ABC endurance score: F=2.62 (1, 4), p<0.01]. In contrast the interaction trotting style*speed produced significant lower scores in the 0-5 LGS (b=0.25, p<0.01) and ABC endurance scores (b=0.23, p=0.02). The mixed model analysis indicated a significant effect of the trotting style, and handler on both subjective results, as depicted in table 3.

Table 3: Fixed effects of trotting style, speed, the interaction between trotting style and speed, and handler on the 0-5 lameness grading scale (LGS) and the ABC endurance score.

	0-5 LGS				**ABC SCORE**			
	F	df1	df2	p-value	*F*	df1	df2	p-value
Trot Style	5,502	1	26	**0,027**	13,235	1	406	**0,000**
Speed	1,867	1	29	0,182	0,790	1	360	0,375
Trot Style * Speed	6,661	1	22	**0,017**	13,682	1	406	**0,000**
Handler	3,008	10	28	**0,010**	7,643	10	7	**0,008**

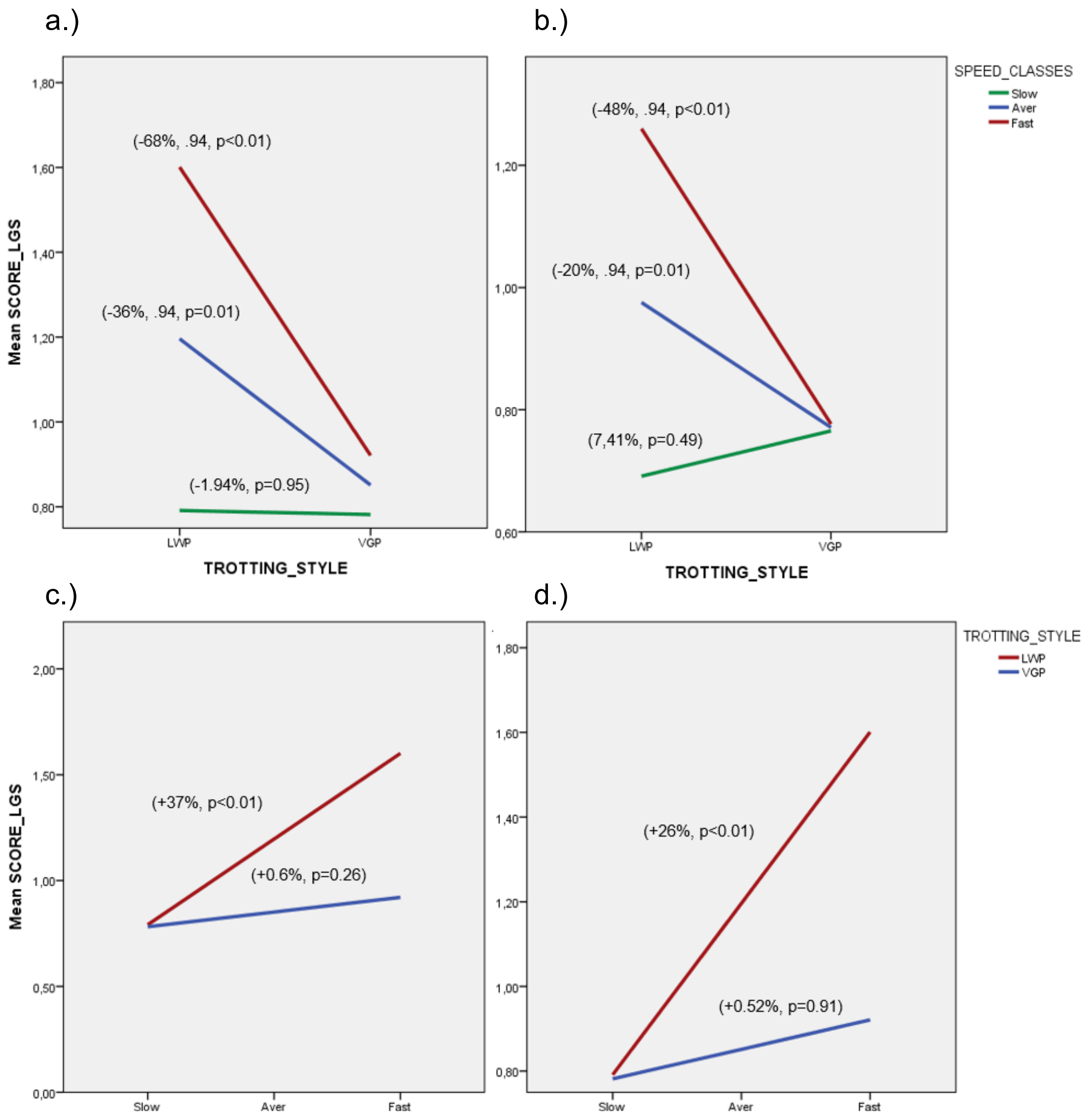


Fig. 7: Moderator effects of low (calculated as 1 SD above the mean=10.74 km/h), average and high speed (calculated as 1 SD above the mean=11.83 km/h) on lameness work-up (LWP) and vet gate presentation (VGP) styles in 0-5 LGS (a.) and ABC endurance scores (b.). Moderation effects of trotting and effects of trotting styles LWP and VGP on speed classes in0-5 LGS (c.) and ABC endurance scores (d.). A, B and C were coded into 0, 1 and 2 respectively.

Discussion

This study investigated the agreement between six FEI 4-star endurance official veterinarians recruited worldwide in grading lameness using a numerical 0-5 lameness grading scale and the A, B and C score used during endurance competitions on slightly lame horses as defined by their owner/handler/vet, actively competing in endurance, while trotted by experienced handlers in two presentation styles (LWP/VGP). The hypotheses that a VGP would generate lower scores and would be performed at a higher speed than a LWP were accepted. However, the higher speed in the VGP seems to be unrelated to the lower scores assigned.

Subjective evaluation

This study showed that veterinarians might have a different understanding of the ABC endurance score when classifying lameness in a 0 to 5 scale, which might reflect the individual and regional interpretation of such scores in the absence of FEI clearer guidelines. Raters in this study were indeed from countries across four different continents, all of which have their own national rules, where ‘fit to compete’ and ‘fail to qualify’ gaits are explained in more detail, which they might use interchangeably in national and FEI competitions. Since many practitioners are accustomed to the AAEP scale to grade lameness worldwide, we decided to use a scale with the same order of magnitude, e.g., 0 to 5 in order to investigate the ABC endurance score.

As opposed to the AERC guidelines, where AAEP grades 1 (lameness is difficult to observe and is not consistently apparent, regardless of circumstances) and 2 (lameness is difficult to observe at a walk or when trotting in a straight line but consistently apparent under certain circumstances) are acceptable to compete (AERC, 2016), our study suggests a threshold of grade 1 for fitness to compete. Therefore, the raters participating in this study seemed to be more conservative than the AERC guidelines, indicating that a grade 2 out of 5 was a fail to qualify or endurance C score, and not a grade 3, defined by the AAEP as a lameness consistently observable at a trot under all circumstances. This result could help setting more detailed guidelines for the endurance score and standardise lameness scoring criteria.

Objective gait assessment

Our results showed a disagreement between objective and subjective evaluations with 25% of the trot-ups deemed fit to compete by veterinarians showing a severe gait asymmetry (>18mm in the front or 9mm in the hindlimb) with the IMUs gait analysis. An experimental study performed during an endurance competition also found disagreement between the objective measurements and subjective ABC endurance score, however, the disagreement disappeared when horses found to be mildly lame were requalified as sound, and only those with moderate to severe lameness were considered lame (C grade) (Lopes *et al.*, 2018). In our study, if we would have considered that threshold, another quarter of the horses deemed to be moderately lame, would have also failed, e.g., half of the horses that passed in the subjective evaluation would have been considered not fit to compete. FEI endurance veterinarians have previously positively commented on the use of a body-mounted IMUs system (Lopes *et al.*, 2018) with two thirds being receptive to have access to an objective system that would help them quantify gait asymmetry and corroborate their decisions (de Mira M.C. *et al.*, 2019). Furthermore, there was no interference of the waiting time necessary to instrument a horse with IMUs on the second heart rate evaluation used to measure the cardiac recovery index in endurance competitions (de Mira *et al.*, 2020). However, the high sensitivity and low specificity of such systems shown by a lack of agreement with the subjective assessment, still preclude its use in competition settings when using the recommended manufacturer's thresholds for lameness detection (Hammarberg *et al.*, 2016a; Keegan *et al.*, 2010; Keegan *et al.*, 1998; McCracken *et al.*, 2012). This lack of alignment between the subjective scores and objective gait analysis was reflected by the poor predictability for fitness to compete or FTQ obtained in the ROC analysis for the OS calculation in our study. A recent study identified a moderate asymmetry in a large proportion of owner-sound horses (Rhodin *et al.*, 2017), suggesting that horses that are deemed sound to compete by their carers could often be lame if evaluated with an IMU system. However, it was previously reported that thresholds for symmetry must be used cautiously since it is unclear how objective asymmetry translates into pain or any musculoskeletal condition (van Weeren *et al.*, 2017). Even if thresholds might be hard to define during competition, objective gait analysis systems could be useful to assess the progression of a gait asymmetry at various moments throughout a competition. Furthermore, a recent study has shown an expectable substantial lower within-horse than between-horse variation (Hardeman *et al.*, 2019), hence different horses under different situations like during competition of different disciplines, might show different thresholds. Nonetheless, other factors must be taken in account in endurance competitions, such as progressive fatigue, reported before to induce gait changes in the horses' trot-ups (Riber *et al.*, 2006a; b).

As reported in other studies (Pfau *et al.*, 2020; Reed *et al.*, 2020), except for bilateral similar-in-magnitude asymmetries, both single limb and multi-limb gait asymmetries can be identified by objective gait analysis. The multi-limb lameness patterns suggested by Reed et al. (2020) and the withers direction as described by Rhodin et al. (2018) further help to untangle the primary lame limb from the compensatory lameness. In our study, the most frequent configuration was primary frontlimb gait asymmetry with a secondary contralateral push-off and impact hindlimb lameness.

Front versus hindlimb lameness' classification

In this study, the frontlimb was more often identified as a cause of gait asymmetry and with higher severity than the hindlimb in both subjective and objective evaluations. This indicated a higher prevalence of forelimb lameness in our sample, which matches a recent study that identified 77% of 351 orthopaedic injuries as occurring in the forelimb of endurance horses.

However, asymmetries were in proportion more often deemed to be sound in the hind than in the front limbs in the subjective than in the objective evaluation. Less sensitivity to hind limb lameness in the subjective evaluation was reported by other studies (Hammarberg *et al.*, 2016b; Keegan *et al.*, 2010). A 25% difference seems to be necessary for the human eye to detect asymmetry in two moving objects simulating hindlimb (Parkes *et al.*, 2009); (Starke and Oosterlinck, 2019) versus 15% in the front limb(Starke and Oosterlinck, 2019).

Agreement between observers and repeatability of scores

Blinded intra-rater gait assessment or repeatability of lameness scoring has been reported to be moderate to poor (Hammarberg *et al.*, 2016b; Serra Bragança *et al.*, 2020). Indeed, agreement between veterinary observers, as determined by ICC scores, was similar to values obtained from racing Thoroughbreds in training when analysed by experienced racetrack veterinarians (Pfau *et al.*, 2020). Similar to other disciplines, tolerance regarding fitness to compete is interpreted differently by trainers and owners, when compared with a purely clinical assessment (Lopes *et al.*, 2018; Pfau *et al.*, 2016; Rhodin *et al.*, 2017; Serra Bragança *et al.*, 2020). In fact, subjectivity of gait assessment is often questioned by riders, trainers and owners in competitions, eliciting frequent complaints and depreciating comments about failing a horse to qualify (de Mira M.C. *et al.*, 2019).

In a recent study (Serra Bragança *et al.*, 2020), Serra Bragança et al. used Fleiss' Kappa to assess the dichotomous score of fit or not fit to compete in sport horses assessed live and by video, showing a poorer inter-rater agreement (ICC: 0.12-0.33) and a higher intra-rater (ICC:0.48) agreement in the video analysis when compared to our study. Fleiss' Kappa and ICC are measures of agreement, which values are equivalent under general conditions, e.g. in ordinal scales (Fleiss and Cohen, 1973). In our study, raters agreed more often in trot-ups when horses were deemed to be sound or when there was a clear lameness (scores of 2 to 3, out of 5). Moreover, the inter-rater agreement for limb identification was better in the VGP than the LWP assessments. Horses were reported to be more often subjectively classified as sound when a mild lameness was present (<1.5 out of 5) and when trot-ups were performed at a higher pace (Keegan *et al.*, 2010; Starke *et al.*, 2013). However, similarly to our study, it has been previously demonstrated that a faster speed might improve the agreement in correctly identifying the lame limb in a more obvious lameness (Peham *et al.*, 2000; Starke *et al.*, 2013). As expected, there was better agreement on the need to call for a panel in VGP in our study, with the most common reason being an inadequate presentation, e.g., interfering with the horse's head position by pulling or lifting, and not lameness. This might have contributed to the better inter-rater agreement in the ABC endurance score for LWP regarding VGP, since often an inadequate trotting presentation will be classified as B. This was to be expected as the ABC endurance score comprises more than a lameness evaluation, including other factors such as impulsion and presentation adequacy (Bennet *et al.*, 2020).

A good agreement was found in successive measurements using objective readings (Keegan *et al.*, 2011). However, the horse's intra-run variability of lameness in repeated trot-ups makes repeatability (Sepulveda Caviedes *et al.*, 2018) in a training or monitoring setting (Hardeman *et al.*, 2019) cumbersome and needs further investigation. Moreover, videoed trot-ups are known to decrease the agreement when compared to live evaluations (Hammarberg *et al.*, 2016a; Leelamankong *et al.*, 2020; Serra Bragança *et al.*, 2020). A further study performed live during competitions could better characterise subjective and objective repeatability.

Effect of trotting style on subjective assessment

To the authors' knowledge, the impact of the trot-up presentation style on subjective lameness evaluation in endurance horses had not been reported before. This study confirmed that veterinarians grade horses as less lame for the VGP coupled with a higher trotting speed than the LWP style trot-ups when using both the 0-5 LGS 0-5 and ABC endurance score. The way a horse is handled during a trot-up has previously been identified by FEI endurance veterinarians as a major and common limitation to lameness identification in competitions (de Mira M.C. *et al.*, 2019). Most professional handlers seem to intuitively trot up horses differently in a vet gate than for a lameness workup. All handlers in this study were experienced in endurance and promptly acknowledged the difference between the two trotting presentation styles and no further instructions were given to avoid undesired and biased changes. "Legal" techniques commonly used to minimize lameness identification include giving the horse more impulsion, attained by a "cheer up" and a faster pace, but avoiding that a horse falls slowly into walk when lame steps could be more obvious caused by the muscular and joint strain necessary to dampen the movement inertia. Even if FEI rules clearly instruct handlers to trot on a loose lead, other techniques would be pulling and holding the lead rope short or lifting the head to avoid a head nod, mostly when the horse trots away so the view is obscured to the evaluator. Furthermore, the handlers could exploit the previous definition of lameness in FEI rules as an irregular gait consistently observable at the trot, by rendering any perceived lameness as inconsistent as possible, for example by jerking the lead rope. A handler effect was registered in our study. Left or right handler position in relation to the horse was considered before to be negligible on limb loading measured by a force plate, provided the handler did not interfere (Van de Water *et al.*, 2016). In our study, the handler was specifically asked to interfere and to do what they would normally do in a VGP.

Effect of trotting speed on subjective and objective evaluation

This study confirmed the anecdotal impression that horses are trotted at a faster pace in the vet gate when compared to a regular lameness evaluation. Speed was previously reported to influence mild lameness subjective identification (Keegan *et al.*, 2010). Starke et al (2013) reported that there were more mildly lame horses declared sound in a straight line at higher speeds by veterinarians in a video-analysis study. Speeds were determined by a sole handler instructed to trot a slow, preferred and fast pace (Starke *et al.*, 2013). In our study, a faster speed coupled with the VGP style was a determinant factor for the lower scores. Speed or a trotting style on their own seem not to be decisive for the subjective classification. This could be because multiple handlers were used with different running abilities. Furthermore, our handlers were adapting the trot-up to a vet gate presentation as conceived by them, and not just running faster. The lesser effect of speed in scores could also be explained by the fact that raters did not just evaluate the strides that matched the objective analysis, as in Starke et al.'s study (Starke *et al.*, 2013), but the complete trot-up. Also, veterinarians were exposed in our study to additional factors other than speed, such as acceleration and deceleration during the turn, more accurately reflecting a real setting such as competition, which could have influenced their decision-making. Conversely, speed can increase the likelihood of a horse being declared sound in subtle lameness (Starke *et al.*, 2013), yet the inter-rater agreement in identifying the correct lame limb in more obvious lameness was reported to improve at higher speeds (Peham *et al.*, 2000). However, since speed can either limit or enhance lameness according to its severity, the effect of speed could have been different if analysed in different lameness categories.

It was recommended by another study that trotting speed should be consistent between trot-ups to minimise the influence of the baseline lameness degree when using gait analysis systems, but this seems to be more important during trot in the circle (Starke *et al.*, 2013). In our case the overall symmetry (OS) value was not significantly affected by the trotting speed or trotting style. Since the subjective evaluation in our study seemed to be particularly affected by the trotting style with faster speeds but not at slow speed, it should be kept slow to moderate between trot-ups, including those performed for the voting panel. It might be recommendable to introduce more specific guidelines for the handler in the rules and education resources and avoid handler's change, as often happens when the voting panel is called.

Moderation effect of speed on the presentation trotting style

Trotting style and speed did not show a significant effect on subjective scores unless the interaction was introduced in the model. A moderator is described as a third variable that changes the association between an independent and an outcome variable. The explanatory variables have not to be significant predictors for the outcome but still have an effect as an interaction. Failing to consider its effect in the data may lead to an erroneous explanation for an outcome to be missed (Bennett, 2000). In our case trotting styles had a different outcome depending on speed, e.g., when speed was low no impact or difference was seen on outcome, increasing proportionally to speed. Interestingly speed elicited higher scores when horses were trotted in the LWP style. Therefore, speed on its own seemed to worsen gait scores provided horses are trotted loose.

Limitations

The current study did have several limitations. Video-analysis does not entirely reflect live evaluations and is known to generate a lower agreement between raters (Hammarberg *et al.*, 2016b; Keegan *et al.*, 2010; Leelamankong *et al.*, 2020; Serra Bragança *et al.*, 2020). The videos' order was randomised when presented to veterinarians, but the selection was according to quality and lame limb distribution. On the other hand, handlers order was randomised, but not the presentation styles order performed by each handler. Also, some horses might have warmed up or deteriorated during the requested 12 trot-ups. The trotting surface effect was assessed grouping the horses in each country due to trotting lane similarity e.g., a built-in hard versus natural dirt trotting lanes, but 'country' may have had other subtle differences not accounted for here. The interaction of handler with trotting style could not be analysed due to study design, e.g., inequity and size of the sample. The sample size was considered adequate for this study, however the threshold found for the 0-5 LGS might not represent the population of FEI registered horses worldwide. Our results should be used with caution when extrapolating to other populations. The categorization used to classify objectively gait asymmetry, was created theoretically in another study (Hammarberg *et al.*, 2016a) and it is not known if it could differ from the 0 to 5 lameness grading scale used. How front and hindlimb objective asymmetry scores relate or differ in severity regarding subjective grading needs further investigation.

Conclusions:

Since gait subjective grading showed a poor agreement between raters and seemed to be significantly affected by trot-up presentation style and higher trotting speed, uniformisation of horses' handling should be attempted in competitions, bringing more accuracy to veterinarians' evaluation and fair-play between competitors. Moreover, veterinarians seem to have a different interpretation of the ABC endurance score in relation to lameness grades in a 0-5 scale. A cut-off value for lameness >1 out 5 is proposed by our results to be indicative of absence of fitness to compete, but further studies might be needed to support the inclusion of this threshold to distinguish between a B and a C trot-up.

The IMU objective analysis is more sensitive than the subjective assessment on scoring lameness, with one quarter of the trot-ups deemed to be subjectively fit to compete, identified as severe gait asymmetry. Moreover, lower scores were given more often in the hind limb when compared with front limb lameness, even though the objective scores did not reflect such a difference. Objective gait analysis studies in real settings could accelerate its calibration for its inevitable utilization in competitions in a near future.

Guidelines and further educational resources could help to uniformise FEI endurance veterinarian's gait assessment criteria and benefit the endurance community. This would be especially important in larger events, such as world championships, where veterinarians from different regions come together and whose criteria might differ. Further trot-up guidelines to handlers could also facilitate a more consistent gait assessment by veterinarians.

Manufacturers' addresses:

[a] Lameness Locator ™, Equinosis LLC, Columbia, Missouri, USA

[b] Pixio by Move'n See, Brest, France

[c] Polar GPS m400, Inc. Kempele, Finland)

[d] SPSS, IBM, Armonk, New York, USA

References:

AERA, 2016. Rulebook - Section 3 - Veterinary Rules. Available at: http://aera.asn.au/national-rules/.

AERA, 2020. Guide to Australian Endurance Vetting. Available at: https://aera.asn.au/guide-australian-endurance-vetting/.

AERC, 2016. Guidelines for Control Judges and Treatment Veterinarians at AERC Endurance Competitions. Available at: https://aerc.org/static/upload/CJHandbook2016.pdf.

Bennet, E.D., Hayes, M.E., Friend, L. and Parkin, T.D.H., 2020. The association between clinical parameters recorded at vet gates during Fédération Equestre Internationale endurance rides and the imminent risk of elimination. Equine Vet J 52: 832-840. https://doi.org/10.1111/evj.13264

Bennett, J.A., 2000. Mediator and moderator variables in nursing research: Conceptual and statistical differences. Research in Nursing & Health 23: 415-420. https://doi.org/10.1002/1098-240X(200010)23:5<415::AID-NUR8>3.0.CO;2-H

de Mira M.C., Santos, C., Lopes, M.A. and Marlin, D.J., 2019. Challenges encountered by Federation Equestre Internationale (FEI) veterinarians in gait evaluation during FEI endurance competitions: an international survey. Comparative Exercise Physiology 15: 371-378. https://doi.org/10.3920/cep180058

de Mira, M.C., Williams, J., Santos, C., Rodrigues, P., Arroja, B. and Marlin, D., 2020. Do waiting times in endurance vet gates affect the cardiac recovery index? Comparative Exercise Physiology: 1-8. https://doi.org/10.3920/CEP190081

ERASA, 2018. Veterinary Rules. Available at: https://www.erasa.co.za/library/front/lib_templates/2001/files/Chapter%208%20JULY%202018%20v2.pdf.

FEI, 2013. Endurance Pilot Global Endurance Injuries Study (GEIS). Fédération Equestre Nationale, Lausanne, Switzerland.

FEI, 2021a. Endurance Rules. Available at: https://inside.fei.org/sites/default/files/FEI%20Endurance%20Rules%20-%201%20January%202021%20-%20CLEAN.final_.pdf.

FEI, 2021b. Veterinary Regulations. Available at: https://inside.fei.org/sites/default/files/2021%20VRs%20-%20Clean.pdf.

Fleiss, J.L. and Cohen, J., 1973. The Equivalence of Weighted Kappa and the Intraclass Correlation Coefficient as Measures of Reliability. Educational and Psychological Measurement 33: 613-619. 10.1177/001316447303300309

Griffin, T.M., Kram, R., Wickler, S.J. and Hoyt, D.F., 2004. Biomechanical and energetic determinants of the walk–trot transition in horses. Journal of Experimental Biology 207: 4215-4223. 10.1242/jeb.01277

Hammarberg, M., Egenvall, A., Pfau, T. and Rhodin, M., 2016a. Rater agreement of visual lameness assessment in horses during lungeing. Equine Vet J 48: 78-82. 10.1111/evj.12385

Hammarberg, M., Egenvall, A., Pfau, T. and Rhodin, M., 2016b. Rater agreement of visual lameness assessment in horses during lungeing. Equine Vet J 48: 78-82. 10.1111/evj.12385

Hardeman, A.M., Serra Bragança, F.M., Swagemakers, J.H., van Weeren, P.R. and Roepstorff, L., 2019. Variation in gait parameters used for objective lameness assessment in sound horses at the trot on the straight line and the lunge. Equine Vet J 51: 831-839. https://doi.org/10.1111/evj.13075

Keegan, K.G., Dent, E.V., Wilson, D.A., Janicek, J., Kramer, J., Lacarrubba, A., Walsh, D.M., Cassells, M.W., Esther, T.M., Schiltz, P., Frees, K.E., Wilhite, C.L., Clark, J.M., Pollitt, C.C., Shaw, R. and Norris, T., 2010. Repeatability of subjective evaluation of lameness in horses. Equine Vet J 42: 92-97. https://doi.org/10.2746/042516409X479568

Keegan, K.G., Kramer, J., Yonezawa, Y., Maki, H., Pai, P.F., Dent, E.V., Kellerman, T.E., Wilson, D.A. and Reed, S.K., 2011. Assessment of repeatability of a wireless, inertial sensor-based lameness evaluation system for horses. Am J Vet Res 72: 1156-1163. 10.2460/ajvr.72.9.1156

Keegan, K.G., Wilson, D.A., Wilson, D.J., Smith, B., Gaughan, E.M., Pleasant, R.S., Lillich, J.D., Kramer, J., Howard, R.D., Bacon-Miller, C., Davis, E.G., May, K.A., Cheramie, H.S., Valentino, W.L. and van Harreveld, P.D., 1998. Evaluation of mild lameness in horses trotting on a treadmill by clinicians and interns or residents and correlation of their assessments with kinematic gait analysis. Am J Vet Res 59: 1370-1377.

Koo, T.K. and Li, M.Y., 2016. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med 15: 155-163. 10.1016/j.jcm.2016.02.012

Leelamankong, P., Estrada, R., Mählmann, K., Rungsri, P. and Lischer, C., 2020. Agreement among equine veterinarians and between equine veterinarians and inertial sensor system during clinical examination of hindlimb lameness in horses. Equine Vet J 52: 326-331. 10.1111/evj.13144

Lopes, M.A.F., Eleuterio, A. and Mira, M.C., 2018. Objective Detection and Quantification of Irregular Gait With a Portable Inertial Sensor-Based System in Horses During an Endurance Race—a Preliminary Assessment. Journal of Equine Veterinary Science 70: 123-129. https://doi.org/10.1016/j.jevs.2018.08.008

Marlin, D. and Williams, J., 2018. Equine endurance race pacing strategy differs between finishers and non-finishers in 120 km single-day races. Comparative Exercise Physiology 14: 11-18. https://doi.org/10.3920/cep170027

McCracken, M.J., Kramer, J., Keegan, K.G., Lopes, M., Wilson, D.A., Reed, S.K., LaCarrubba, A. and Rasch, M., 2012. Comparison of an inertial sensor system of lameness quantification with subjective lameness evaluation. Equine Vet J 44: 652-656. 10.1111/j.2042-3306.2012.00571.x

Parkes, R.S., Weller, R., Groth, A.M., May, S. and Pfau, T., 2009. Evidence of the development of 'domain-restricted' expertise in the recognition of asymmetric motion characteristics of hindlimb lameness in the horse. Equine Vet J 41: 112-117.

Peham, C., Licka, T., Mayr, A. and Scheidl, M., 2000. Individual speed dependency of forelimb lameness in trotting horses. Vet J 160: 135-138. 10.1053/tvjl.2000.0483

Persson-Sjodin, E., Hernlund, E., Pfau, T., Haubro Andersen, P., Holm Forsström, K. and Rhodin, M., 2019. Effect of meloxicam treatment on movement asymmetry in riding horses in training. PLoS One 14: e0221117. 10.1371/journal.pone.0221117

Pfau, T., Noordwijk, K., Sepulveda Caviedes, M.F., Persson-Sjodin, E., Barstow, A., Forbes, B. and Rhodin, M., 2018. Head, withers and pelvic movement asymmetry and their relative timing in trot in racing Thoroughbreds in training. Equine Vet J 50: 117-124. https://doi.org/10.1111/evj.12705

Pfau, T., Parkes, R.S., Burden, E.R., Bell, N., Fairhurst, H. and Witte, T.H., 2016. Movement asymmetry in working polo horses. Equine Vet J 48: 517-522. 10.1111/evj.12467

Pfau, T., Sepulveda Caviedes, M.F., McCarthy, R., Cheetham, L., Forbes, B. and Rhodin, M., 2020. Comparison of visual lameness scores to gait asymmetry in racing Thoroughbreds during trot in-hand. Equine Veterinary Education 32: 191-198. 10.1111/eve.12914

Reed, S.K., Kramer, J., Thombs, L., Pitts, J.B., Wilson, D.A. and Keegan, K.G., 2020. Comparison of results for body-mounted inertial sensor assessment with final lameness determination in 1,224 equids. J Am Vet Med Assoc 256: 590-599. 10.2460/javma.256.5.590

Rhodin, M., Egenvall, A., Haubro Andersen, P. and Pfau, T., 2017. Head and pelvic movement asymmetries at trot in riding horses in training and perceived as free from lameness by the owner. PLoS One 12: e0176253. 10.1371/journal.pone.0176253

Rhodin, M., Persson-Sjodin, E., Egenvall, A., Serra Braganca, F.M., Pfau, T., Roepstorff, L., Weishaupt, M.A., Thomsen, M.H., van Weeren, P.R. and Hernlund, E., 2018. Vertical movement symmetry of the withers in horses with induced forelimb and hindlimb lameness at trot. Equine Vet J 50: 818-824. 10.1111/evj.12844

Riber, C., Cuesta, I., Munoz, A., Gata, J., Trigo, P. and Castejon, F.M., 2006a. Equine locomotor analysis on vet-gates in endurance events. Equine Vet J Suppl 36: 55-59. 10.1111/j.2042-3306.2006.tb05513.x

Riber, C., Cuesta, I., Munoz, A., Gata, J., Trigo, P. and Castejon, F.M., 2006b. Equine locomotor analysis on vet-gates in endurance events. Equine Vet J Suppl: 55-59.

Sepulveda Caviedes, M.F., Forbes, B.S. and Pfau, T., 2018. Repeatability of gait analysis measurements in Thoroughbreds in training. Equine Vet J 50: 513-518. https://doi.org/10.1111/evj.12802

Serra Bragança, F.M., Brommer, H., van den Belt, A.J.M., Maree, J.T.M., van Weeren, P.R. and Sloet van Oldruitenborgh-Oosterbaan, M.M., 2020. Subjective and objective evaluations of horses for fit-to-compete or unfit-to-compete judgement. The Veterinary Journal 257: 105454. https://doi.org/10.1016/j.tvjl.2020.105454

Starke, S.D. and Oosterlinck, M., 2019. Reliability of equine visual lameness classification as a function of expertise, lameness severity and rater confidence. Veterinary Record 184: 63-63. 10.1136/vr.105058

Starke, S.D., Raistrick, K.J., May, S.A. and Pfau, T., 2013. The effect of trotting speed on the evaluation of subtle lameness in horses. Vet J 197: 245-252. https://doi.org/10.1016/j.tvjl.2013.03.006

Van de Water, E., Oosterlinck, M. and Pille, F., 2016. The effect of perineural anaesthesia and handler position on limb loading and hoof balance of the vertical ground reaction force in sound horses. Equine Vet J 48: 608-612. https://doi.org/10.1111/evj.12491

van Weeren, P.R., Pfau, T., Rhodin, M., Roepstorff, L., Serra Braganca, F. and Weishaupt, M.A., 2017. Do we have to redefine lameness in the era of quantitative gait analysis? Equine Vet J 49: 567-569. 10.1111/evj.12715

Supplementary Material

Figure S1: Gait assessment in endurance horses' quiz

We kindly ask you to replay the videos no more THREE TIMES as it is the standard practice in competitions.

1. **After viewing this video, please assign a score to gait, using a lameness scale from 0 (no lameness visible) to 5 (nonweight-bearing lame).**

a. Grade 0
b. Grade 1
c. Grade 2
d. Grade 3
e. Grade 4
f. Grade 5

2. **Using the FEI endurance scale, how would you grade this trot-up?**

a. A
b. B
c. C

3. **Would you call for a panel?**

a. No
b. Yes, because the horse is lame
c. Yes, because of presentation

4. **In this trot-up you see a:**

a. LF lameness
b. RF lameness
c. LH lameness
d. RH lameness
e. a multiple lameness

5. **If you see a multiple lameness, what is your best guess**
   a. It's most likely a compensatory lameness
   b. It's most likely a true multiple lameness with more than one limb affected
   c. Can't tell

6. **If you think there is a compensatory (false) lameness, which limb do you think it's the source of the primary (true) lameness?**
   a. LF
   b. RF
   c. LH
   d. RH
   e. Can't tell

7. **If you think there is a true multiple lameness, which limbs seem to you to be affected (more than answer allowed).**
   a. LF
   b. RF
   c. LH
   d. RH

Table S1: Guidelines of different governing bodies for lameness evaluation at endurance competitions, AERA (Australian Endurance Riding Conference), AERC (American Endurance Riding Conference), ERASA (Endurance Ride Association of South Africa).

	Fit to compete	Fail to qualify	Further guidance
AERA (2016)	**A**= Willing, strong, normal **B**= Subtle reluctance **C**= Reluctance, tired, not consistently lame	**D=** Unwilling, no animation, consistently lame Suggested criteria for elimination: - Consistent irregularity to and from observer - Able to identify limb i.e. LF or RH (although not necessary) Displays an irregularity of gait exhibited as, but not limited to, a consistent, head bob, hip hike or shortened stride or, an irregularity of gait that threatens the immediate ability of the horse to safely perform athletically.	Any unusual feature about a horse's gait that does not remove a horse from the ride should be noted on the logbook or vet card, especially when detected at the pre-ride inspection. This allows veterinarians at subsequent inspections to be informed of the horse's earlier gait and thus be in a better position to make a judgement on the current gait, as any action taken in respect to a gait abnormality and/ or injury is determined by any deterioration or improvement that has occurred since the previous inspection. If any noted gait abnormality or injury has not deteriorated, the horse will be deemed fit to continue in the ride.
AERC (2016)	**Grade I**= Difficult to observe. Not consistently apparent regardless of circumstances(i.e., weight carrying, circling, inclines, hard surface, etc.). **Grade II**= Difficult to observe at a walk or trotting a straight line; consistently apparent under certain circumstances (i.e., weight carrying, circling, inclines, hard surface, etc.).	**Grade III**= Consistently observable at a trot under all circumstances. At endurance competitions, this generally means consistently observable in both directions of a straight out-and-back trot. (Note: as a general rule of thumb, "consistent" can be defined as observable more than 70% of the time.)	Time pressures require judgments to be rapid and critical. Grades I and II can usually "go on" with caution. As in the pre-ride exam, consider the prognosis of the lameness with work. If the athletic future of the horse is threatened then consider disqualifying the horse.
ERASA ((2018)	**A=** Sound, no signs of unevenness or lameness **B=**Uneven, but not consistently lame	**C=**Lame	
FEI (FEI 2020b)	**A=** Willing, strong, normal(FEI 2013) **B=** Reluctance, tired, not consistently lame-elimination??(FEI 2013)	**C=** Unwilling, no animation, consistently lame-elimination!!! (FEI 2013)	Guidance and definitions from a course provided for FEI1* endurance vets (Drs Martha Misheff and Sarah Coombs personal communication). Note that these definitions come from experienced endurance vets and are not currently included in the FEI endurance rules. (Bennet et al. 2020) **A=** Is ok. We don't say 'sound' because that suggests perfect, but on the basis of 'rhythm, elasticity and vigour' - and I would add symmetry- they don't give cause for concern and are fit to continue.(Bennet and Parkin 2020) **B=** Is passable but not as good, maybe the odd inconsistent stride or needs to be trotted for a second opinion but passed. Often can be indicative of a tired/slightly dehydrated or electro. depleted horse not showing itself off very well. **C=** Is clearly lame

AERA. (2016). Rulebook - Section 3 - Veterinary Rules. Efective January 1st 2016. Retrieved 17.11, 2020, from http://aera.asn.au/national-rules/.

AERC. (2016, April, 2016). Guidelines for Control Judges and Treatment Veterinarians at AERC Endurance Competitions. 3.0 Rev. Retrieved 17.11, 2020, from https://aerc.org/static/upload/CJHandbook2016.pdf.

Bennet, E. D., M. E. Hayes, L. Friend and T. D. H. Parkin (2020). The association between clinical parameters recorded at vet gates during Fédération Equestre Internationale endurance rides and the imminent risk of elimination. Equine Veterinary Journal, DOI: https://doi.org/10.1111/evj.13264, **52**(6): 832-840.

Bennet, E. D. and T. D. H. Parkin (2020). The impact of the mandatory rest period in Fédération Equestre Internationale endurance events. Equine Veterinary Journal, DOI: https://doi.org/10.1111/evj.13148, **52**(2): 268-272.

ERASA. (2018, 8th September). Veterinary Rules.   Retrieved 17.11, 2020, from https://www.erasa.co.za/library/front/lib_templates/2001/files/Chapter%208%20JULY%202018%20v2.pdf.

FEI (2013). Endurance Pilot Global Endurance Injuries Study (GEIS). Lausanne, Switzerland, Fédération Equestre Nationale.

FEI. (2020b, 1 Jan 2020). Veterinary Regulations. 14th Edition. Retrieved 18.11.2020, 2020, from https://inside.fei.org/fei/regulations/veterinary.

Table S2: a.) Used guidelines to classify gait symmetry according to Lameness Locator® Instruction Manual and Reed et al. (2020).
a.)

FRONT LIMB

	MaxDiff > +|6mm| POS	MaxDiff close to zero	MaxDiff > +|6mm| NEG -
MinDiff >|6mm| POS +	IMPACT RF	IMPACT (midstance) RF	PUSHOFF RF
MinDiff close to zero	PUSHOFF RF		PUSHOFF LF
MinDiff >|6mm| NEG -	PUSHOFF LF	IMPACT (midstance) RF	IMPACT LF

HIND LIMB

	MaxDiff > +|3mm| POS+	MaxDiff close to zero	MaxDiff > +|3mm| NEG-
MinDiff >|3mm| POS+	IMPACT AND PUSHOFF RH	IMPACT RH	IMPACT RH PUSHOFF LH
MinDiff close to zero	PUSHOFF RH		PUSHOFF RH
MinDiff >|3mm| NEG	IMPACT LH PUSHOFF RH	IMPACT LH	IMPACT AND PUSHOFF LH

- FRONT LIMB lameness: when mean head vector sum (square root of [$HDmax^2$ + $HDmin^2$) > 8.5 mm, and the SD of HDmin < 120%. MaxDiff.
- LEFT LIMB= negative value; RIGHT LIMB=positive value
- IMPACT 1= and with the same sign direction (+,+) or (-,-)
- IMPACT 2 (midstance)= MinDiff >|6mm| and
- PUSHOFF 1: MinDiff >|6mm| and MaxDiff > |6mm| with opposite signs (+,-) or (-,+), MinDiff dictates the direction of the lameness
- PUSHOFF 2: MaxDiff > |6mm|, difficult to interpret

b.) Distribution (n) of the objective readings obtained by our study according to multiple limb asymmetry patterns as described by Reed et al. (2020) in a population of 1224 equids (numbers in brackets represent the ranking order), e.g., **Patterns 1.1** (1st) **and 1.2** (3)**-** First law of sides; **Pattern 1.3** (4)**-** 45% front or 34% hind or in 11% both limbs; **Patterns 2.1** (2) **and 2.2** (5) **-** Second law of sides. **Pattern 2.3** (6) **-** 2nd law of sides (75% front and 6% hind or 19% both). **3.1** (8)**-** 20% front +35 hind +20% both limbs**. Pattern 3.2** (7)**-** 64% front limb, and withers signal interpretation as reported by Rhodin et al. (2018). LF (left front). RF (right front). LH (left hind). RH (right hind).

	LF and WD MinDiff POS +	**LF and WD MinDiff NEG -**	**RF and WD MinDiff POS +**	**RF and WD MinDiff NEG -**
LH IMPACT + PUSHOFF	Pattern 1.1 LF + LH	Pattern 1.1 LF + LH	Pattern 2.1 RF + LH	Pattern 2.1 RF + LH
LH PUSHOFF	Pattern 1.2 LF + LH	Pattern 1.2 (n=2) LF + LH	Pattern 2.2 RF + LH	Pattern 2.2 RF + LH
LH IMPACT	Pattern 1.3LF + LH	Pattern 1.3 LF + LH	Pattern 2.3 RF + LH	Pattern 2.3 RF + LH
RH IMPACT + PUSHOFF	Pattern 2.1 LF + RH	Pattern 2.1 LF + RH	Pattern 1.1 RF + RH	Pattern 1.1 RF + RH
RH PUSHOFF	Pattern 2.2 LF + RH	Pattern 2.2 LF + RH	Pattern 1.2 (n=1)RF + RH	Pattern 1.2 RF + RH
RH IMPACT	Pattern 2.3 LF + RH	Pattern 2.3 LF + RH	Pattern 1.3 RF + RH	Pattern 1.3 RF + RH
LH PUSHOFF + RH IMPACT	Pattern 3.1 LF + LH + RH	Pattern 3.1 LF + LH + RH	Pattern 3.2 RF + RH + LH	Pattern 3.2 RF + RH + LH
LH IMPACT + RH PUSHOFF	Pattern 3.2 LF + LH + RH	Pattern 3.2 LF + LH + RH	Pattern 3.1 RF+RH+LH	Pattern 3.1 RF+RH+LH

[1]Reed, S.K., Kramer, J., Thombs, L., Pitts, J.B., Wilson, D.A. and Keegan, K.G. (2020) Comparison of results for body-mounted inertial sensor assessment with final lameness determination in 1,224 equids. *Journal of the American Veterinary Medical Association* **256**, 590-599.

[2]Rhodin, M., Persson-Sjodin, E., Egenvall, A., Serra Braganca, F.M., Pfau, T., Roepstorff, L., Weishaupt, M.A., Thomsen, M.H., van Weeren, P.R. and Hernlund, E. (2018) Vertical movement symmetry of the withers in horses with induced forelimb and hindlimb lameness at trot. *Equine veterinary journal* **50**, 818-824.

Table S3: Subjective grading of videoed trot-ups' performed by six 4*FEI (with known identification), e.g. ABC endurance scores and 0-5 lameness grading scale, and limb identification, aligned with the objective measurement results (degree and asymmetric limbs with the primary limb written in bold). Repeated videos for subjective evaluations are written in red. N represents the number of subjective evaluations, including the questionnaires with unknown identification. Medians of 0-5 lameness grading scale (M0-5) and of ABC endurance scores (Mend) are presented. LWP (lameness work-up presentation). VGP (vet gate presentation). NL (non-lame), LF (left front), RF (right front), LH (left hind), RH (right hind).The primary gait asymmetry is written in bold. Whenever veterinarians identified more than one limb lameness, values are shown between brackets.

	Trot Style	H TR	Vet 1		Vet 2		Vet 3		Vet 4		Vet 5		Vet 6		N VETS	M0-5	MEnd	OBJECTIVE RESULTS	INTERPRETATION
Horse 1	**LWP**	A 1	0	A		A	1 RF	B	1 RF	C	0	A				0	A	Sound	
		B 1	0	A	RF	B	1	B								1	B	Sound	
		C 2	1 LF	B		A	0	A								0	A	Sound	
	VGP	A 1	0	A		A	0	A	0	A	0	A	0 0	A		0	A	Sound	
		B 2	0	A		B	1 RF	B								1	B	Sound	
		C 2	1	B												1	B	Sound	
Horse 2	**LWP**	A 1	3 RH	B	RF	C	3 RF	C	2 RF	C	2 RF	B	2 RF	C		2	C	Severe	RF+LH
	REP		3 RH	C	RF	C	2 RF	C	1 RF	C	2 RF	C							
		B 2	3 LH	C	LF	C	1 RH	B			2 RF	B	3 RF	C		3	C	Severe	RF+LH
		C 1	2 LH	B	RF	B	1 RF	B								2	C	Severe	RF+LH
	VGP	B 2	0	A	RF	C	0	A	2 RH	C	2 RF	B				2	C	Moderate	RF
	REP		3 LH	C	RF	C	1 RF	B	2 LH	C	2 RF	B	2 RF	C					
	VGP	C 2	3 LH	C												2	C	Severe	RF+LH
Horse 4	**LWP**	A 1	3 RH	B			0	A	2 RH	C	1 RH	B	3 LH	C		2	B	Sound	
		B 1	2 RH	B		A										2	B	Severe	LH
		C 1	2 RH	B		A	0	A								1	B	Mild	LH
	VGP	A 2	3 RH	B	LF	B	0	A	2 RH	B	1 RH	B	0 0	B		1	B	Sound	
	VGP	B 2	2 RH	B	LF	B	0	A								1	B	Mild	LH
	REP		2 LH	B	MULT	B	0	A	2 RH	C	1 RH	B							
	VGP	C 1	1	B	RF	B	1	B			2 RH	B				1	B	Sound	
	REP			B		A	0 LF-RH	A	1 LF	C									
Horse 5	**LWP**	A 1	3 LF	B	RF	B	2 LF	C	3 LH	C	1 LF	B	2 LF	C		2	C	Severe	LF
		B 2	1	B		A										0	B	Moderate	LF
		C 1	2 MULT	C	RF	B	1 LH	B	3 LH	C	0	0 A	2 LH	C		1	B	Moderate	LF
	REP		2	B			0	A	2 RF-LH	C	0 LF-RH	A							
	VGP	A 1	3 LF-RH	C	RF	B										2	C	Moderate	LF
		C 1	1	B			0	A	2 LH	B	1 RH	B	2 LF	C		1	B	Moderate	LF
Horse 6	**LWP**	A 1	0	A		A	0	A								0	A	Mild	RH
		B 2	2 MULT	B		A	2 RH	C	2 LH	C	1 RH	B	2 LH	C		2	C	Moderate	RH
		C 1	0 LF-RF	A		A	0 MULT	B	2 RH	C	1 RH	B	0 0	A		0	B	Mild	RH
	VGP	A 2	0	A		A										0	A	Sound	
			0 RF-LH	A		A	0	A	2 LH (LF-RH)	B	1 RH	B	0 0	A		0	A	Sound	
	REP		2 LF	B		A	1 MULT	B	2 RH	C	0	0 A							
		C 1	1	A			0	A								1	A	Sound	
Horse 8	**LWP**	A 1	1	A		A	0	A								0	A	Sound	
		B 1	1	B		A	0	A	0	A	0	0 A				0	A	Sound	
		C 2	1	B		A	0	A	0	A	0	0 A	0 0	A		1	B	Severe	LF+RH
	VGP	A 1	3 RF	C	LF	B	0	A	1 LF	B	0	0 A				0	A	Moderate	LF
	REP		1	A															
	VGP	B 1	1 LF-RH	B		A	0	A								1	B	Moderate	LF
		C 1	1	B												0	A	Moderate	RH
Horse 9	**LWP**	A 2	2 RF	C		A										0	A	Sound	
		B 2	1	B			0	A								1	B	Sound	
		C 2	1	A	RF	A	0	A	1 RF	B	1 RH	B	2 RH	B		1	B	Mild	RH
	VGP	A 2	0	A		A	0	A								0	A	Moderate	RH
		B 2	3 MULT	C		A	0	A								0	A	Mild	LF
	REP		2 MULT	B		A	0	A	0	A	0	0 A							
	VGP	C 2	2 MULT	B	RF-LH	A	0	A	2 LH	C	0	0 A				0	A	Sound	
Horse 11	**LWP**	A 1	2	B			1 RF	B	3 RF	C	1 LF	B	2 RF	C		1	B	Moderate	RF+LH
	REP		0	A	RF	C	1	B	1 RF	B	0	0 A							
		B 2	3 LH	C	RF	B										1	C	Mild	RF
	VGP	A 2	1	B	RF	C										3	C	Moderate	RF
		B 2	3 RF	C	RF	C	3 RF	C								1	C	Severe	RF
		C 1	3 LF	C			3 RF	C	3 RF	C	2 RF	C	3 RF	C		3	C	Severe	RF
Horse 13	**LWP**	B 1	1	B		A	0	A	0	B	0	0 A	0 0	A		0	A	Sound	
	VGP	B 2	2 LF	C	LF	B	2 LF	C	2 LF	C	0	0 A	0 0	A		2	C	Sound	
		C 2	2	B		A	0	A	0	B	0	0 B	0 0	B		0	B	Sound	
Horse 17	**LWP**	A 2	0	A												0	A	Severe	LF
	VGP	A 2	0	A		A	0	A	0	A	0	0 A				0	A	Sound	
Horse 18	**VGP**	B 3	1	B	LF-RF	C	1 RF	B	1 LF	B	0	0 A	1 MULTI	B		1	B	Mild	LH
Horse 22	**LWP**	A 1	0	A	LF-RH	B	1 LF-RH	B	3 MULTI	C	1 RF-LH	B	2 RF-LH	C		1	B	Moderate	LF+LH
		B 1	0	A												0	A	Moderate	LF+LH
	VGP	A 1	2 LF	B		A	1 RH	B								0	A	Moderate	LH
	REP		0	A	RH	B	0	A	1 RH	C	2 LH	B							
		B 1	2	B	LF	B	1 RH	B	1 LH	B	2 MULTI	C	0 LH	B		1	B	Severe	LH
Horse 23	**LWP**	A 1	1 LF	B	LF	B	0	A								0	B	Sound	
		B 1	1	B		B										0	A	Severe	LF
		C 2	1	A			0	A	0	A	1 LH	B	2 LF	B		1	B	Severe	LF
	VGP	A 1	3 RF	B		A										1	B	Severe	LF
		B 1	1	A			0	A	0	A	1 LH	B	0 0	A		1	B	Severe	LF
		C 1	1	A	LF	C	0	A	0	B	0	0 A	0 0	A		0	A	Severe	LF
Horse 28	**LWP**	C 2	3 MULT	C	RF	C	2 RF(LF-RH)	B	3 RF	C	2 RF	B	0 0	A		2	C	Severe	RF
	VGP	B 2	3 MULT	C												2	C	Severe	RF+RH
	VGP	C 2	0	A	RF	B	1	B	2 LH	C	1 RF	B	1 RF	B		1	B	Severe	RF+RH
Horse 29	**LWP**	A 1	2 RF	C	RF-LH	C	2 LF (RF-LH)	C	3 RH	C	0	0 A	3 MULTI	C		3	C	Severe	RF+LH
		B 1	0	A	MULT	C	1 RF-LH	B	2 RF	C	2 RH	B	2 LH	C		1	B	Moderate	RF+LH
	REP		0	A		C	1	B	1 RF	C	0	0 A							
		C 1	3 RH	B	MULT	C	1 LH	C	2 RF	B	2 RF-LH	C	3 MULTI	C		3	C	Severe	RF+LH
	VGP	A 1	2 RF-LH	B		B	1	B								1	B	Severe	LH
		B 2	1 RF-LH	B												1	B	Severe	RF+LH
		C 2	2	C		A	0	A	0	C	0	0 B				0	B	Moderate	RF+LH
Horse 30	**LWP**	A 1	1	B		A	0	A	1 RF	B	1 RF	B	2 MULTI	C		1	B	Severe	RF+LH
		B 2	1	B		A										1	B	Mild	
		C 2	0	A			0	A	0	A	0	0 A	0 0	A		0	A	Mild	LH
	VGP	A 3	1	B		A	0	A	0	A	0	0 A	2 RF	C		0	A	Severe	RF
	REP		2	B		A	0	A	0	A	0	0 A							
		C 2	0	A		A	0	A	2 RF	B	0	0 A	2 RF	C		0	A	Moderate	LH
Horse 31	**LWP**	B 2	0	A												0	A	Moderate	RH
		C 2	0	A	RF	B	0	A								0	A	Moderate	RF+RH
	VGP	A 1	2	C	RF	B	0	A	0	B	0	0 B	1 0	C		0	B	Sound	
		B 2	2	C			1 RF	B	2 RH	C	1 RF	B	2 RF	C		2	C	Moderate	RF
		C 1	0	A		A										0	A	Sound	
Horse 32	**LWP**	A 1	1	B		A	0	A	0	B	0	0 B	1 RF	B		0	B	Sound	
		A 2	2	B		A	0	A								1	B	Sound	
		B 1	1	B		A	0	A	1 LF	B	0	0 A				0	A	Moderate	RH
		C 2	0	A		A	0	A	0	B	0	0 B	0 0	A		0	A	Moderate	RH
	REP		1	B		B	0	A	0	A	0	0 A							
	VGP	A 2	2	C		A	0	A	0	B	0	0 B	1 RF	A		0	A	Sound	
	REP		1	B		A	0	A		C	0	0 A							
		B 1	2	C		A	0	A		B	0	0 B	1 RF	B		0	B	Mild	RH

Table S4: Objective measurements and interpretation of the results. Results in bold indicate the primary gait asymmetry. Whenever a multiple limb lameness was found, the pattern was indicated as mentioned in Table S2. NL (non-lame), LF (left front), RF (right front), LH (left hind), RH (right hind), I (impact), P (push-off).

Trot_Up_ID	Horse	Handler	Trot Type	Repeated Trot	Stride_Rate_Front	Strides_Front	Ratio_mean_Front	Stride_Rate_Hind	Strides_Hind	Ratio_mean_Hind	Front_Min_Diff	sdHDmin_Front	Variation_Coef_Min_Front	Front_Max_Diff	sdHDmax_Front	Variation_Coef_Max_Front	Mean_Vector_Sum_Front	Hind_Min_Diff	sdHDmin_Hind	Variation_Coef_Min_Hind	Hind_Max_Diff	sdHDmax_Hind	Variation_Coef_Max_Hind	Mean Vector_Sum_Hind	Overall Symmetry	Withers_Min_Diff	sdHDmin_Withers	Withers_Max_Diff	sdHDmax_Withers	Overall Symmetry	Min_W_Signal	Max_W_Signal	FRONT-HIND	Degree	Result	Pattern
1_A_LWP_1	1	A	LWP	1	1,6	14	0,7	1,6	14	0,4	2,7	8,8	325	-4,9	5,6	115	5,6	3,3	6,6	200	-0,4	2,9	808	3,3	6,1	-1,2	8,8	-4,2	8,4	6,1	-	-			Sound	
1_B_LWP_1	1	B	LWP	1	1,6	19	0,8	1,6	19	0,3	10,6	15,0	142	-4,4	10,5	241	11,4	4,2	6,9	166	0,2	5,2	3453	4,2	9,9	3	11,2	-28	20,3	9,9	+	-			Sound	
1_C_LWP_2	1	C	LWP	2	1,5	20	0,6	1,5	20	0,3	5,2	9,1	175	-1,6	7,6	480	5,4	1,5	8,0	543	-1,3	3,6	271	2,0	4,7	5,9	10,9	-9,1	11,3	4,7	+	-			Sound	
1_A_VGP_1	1	A	VGP	1	1,6	17	0,8	1,6	19	0,3	-4,1	7,5	184	1,0	6,2	635	4,2	-3,3	5,6	171	1,2	2,5	215	3,5	5,6	-1,9	11,9	3,9	7,8	5,6	-	+			Sound	
1_B_VGP_2	1	B	VGP	2	1,5	24	0,6	1,5	24	0,3	4,3	10,8	252	-1,4	8,0	564	4,5	4,0	7,1	176	0,9	4,3	478	4,1	6,4	-0,2	13	-21,6	15	6,4	-	-			Sound	
1_C_VGP_2	1	C	VGP	2	1,5	20	1,2	1,5	22	0,2	2,5	9,1	367	4,3	8,0	186	4,9	0,5	4,4	967	0,3	2,5	915	0,5	3,0	-4,1	13,1	-6,6	8,3	3,0	-	-			Sound	
2_A_LWP_1	2	A	LWP	1	1,8	19	1,5	1,8	19	0,4	31,0	16,6	53	19,8	12,4	63	36,8	-9,9	7,4	75	-4,2	5,6	135	10,8	29,2	5,7	8,6	10,7	7,3	29,2	+	+	Multiple	Severe	**RFI**+LHI	2.3
2_B_LWP_2	2	B	LWP	2	1,7	22	1,2	1,7	23	0,4	25,3	16,3	65	16,4	11,5	70	30,1	-7,4	5,5	75	-5,4	4,9	91	9,1	24,2	5,9	9,3	7,6	7,7	24,2	+	+	Multiple	Severe	**RFI**+LHI+P	2.1
2_C_LWP_1	2	C	LWP	1	1,6	25	1,0	1,6	24	0,4	18,3	11,2	61	11,3	9,5	84	21,5	-6,1	5,6	91	-5,2	4,4	83	8,1	18,8	8,5	9,2	7,7	5,7	18,8	+	+	Multiple	Severe	**RFI**+LHI+P	2.1
2_B_VGP_2	2	B	VGP	2	1,7	21	1,1	1,7	21	0,3	17,0	13,7	81	12,5	9,2	73	21,1	-4,4	6,0	138	-3,0	3,5	115	5,3	15,8	10,8	7,8	6,1	6,4	15,8	+	+	FLame Only	Moderate	RFI	
2_C_VGP_2	2	C	VGP	2	1,8	19	1,2	1,8	19	0,4	24,9	12,3	49	19,0	10,7	56	31,3	-4,5	5,7	126	-5,4	5,1	94	7,0	22,7	7,9	10,7	5,5	7,4	22,7	+	+	Multiple	Severe	**RFI**+LHP	2.2
4_A_LWP_1	4	A	LWP	1	1,6	20	0,6	1,6	20	0,3	-1,6	9,9	608	6,1	5,5	90	6,3	-2,4	3,3	138	-1,3	4,2	315	2,7	5,9	21,1	14,4	-0,5	6,5	5,9	+	-				
4_B_LWP_1	4	B	LWP	1	1,6	17	1,1	1,6	18	0,6	-6,7	16,2	244	10,3	19,3	187	12,3	-13,2	8,1	61	-3,3	10,1	304	13,6	19,8	11,3	25	1,5	16,4	19,8	+	+	HLame Only	Severe	LHI	
4_C_LWP_1	4	C	LWP	1	1,5	19	0,8	1,5	19	0,3	-0,2	10,8	5415	1,7	11,5	696	1,7	-4,7	4,6	98	1,3	4,6	363	4,9	5,7	19,9	12,2	-3,6	6,5	5,7	+	-	HLame Only	Mild	**LHI**	
4_A_VGP_2	4	A	VGP	2	1,5	18	0,7	1,5	18	0,4	0,6	6,6	1143	8,8	14,1	159	8,9	-3,4	4,3	126	-2,0	6,3	324	3,9	8,3	19,2	20,4	-5,5	10,4	8,3	+	-				
4_B_VGP_2	4	B	VGP	2	1,5	19	0,6	1,5	20	0,3	-2,0	6,5	322	5,9	8,8	149	6,2	-4,8	3,6	74	-2,0	2,6	133	5,2	8,3	18,9	11,6	0,5	7,1	8,3	+	+	HLame Only	Mild	**LHI**	
4_C_VGP_1	4	C	VGP	1	1,5	19	0,6	1,5	20	0,2	-5,9	6,3	107	1,9	11,4	592	6,2	-2,4	3,3	139	0,1	4,0	8000	2,4	5,5	23,2	12,7	-5,1	6,1	5,5	+	-			Sound	
5_A_LWP_1	5	A	LWP	1	1,6	18	1,0	1,6	20	0,3	-23,3	11,7	50	1,7	8,2	482	23,3	-4,4	8,7	197	0,1	4,5	6457	4,4	16,1	-0,1	9,1	1	8,1	16,1	-	+	FLame Only	Severe	**LFI**	
5_B_LWP_2	5	B	LWP	2	1,6	18	1,1	1,6	19	0,4	-17,8	19,3	109	0,7	10,3	1580	17,8	-1,2	12,4	1082	-0,7	6,6	905	1,4	10,3	-5,8	9	-0,1	10,3	10,3	-	-	FLame Only	Moderate	**LFI**	
5_C_LWP_1	5	C	LWP	1	1,6	22	0,9	1,6	22	0,4	-13,1	10,8	83	3,2	11,6	363	13,5	-5,2	6,4	121	-1,2	5,4	468	5,4	12,1	-0,7	8,4	-1,8	8,6	12,1	-	-	FLame Only	Moderate	**LFI**	
5_A_VGP_1	5	A	VGP	1	1,6	18	1,2	1,6	20	0,4	-14,8	8,0	54	-7,7	12,2	158	16,7	-2,7	9,2	343	4,6	4,7	102	5,4	13,7	-10,3	8,8	-5	10,2	13,7	-	-	FLame Only	Moderate	**LFI**	
5_C_VGP_1	5	C	VGP	1	1,6	19	0,9	1,6	19	0,3	-12,9	14,6	114	5,7	11,3	199	14,0	-2,5	9,9	402	0,4	7,6	1800	2,5	9,5	-7,9	9,2	-0,3	8,9	9,5	-	-	FLame Only	Moderate	**LFI**	
6_A_LWP_1	6	A	LWP	1	1,5	22	0,7	1,5	22	0,3	2,9	8,9	310	-1,0	9,7	964	3,0	-2,4	4,8	204	4,5	1,9	42	5,0	6,6	-2,6	17,5	-7,3	5,1	6,6	-	-	HLame Only	Mild	**RHP**	
6_B_LWP_2	6	B	LWP	2	1,6	18	0,9	1,6	18	0,4	2,0	25,3	1242	-3,9	10,7	273	4,4	1,7	7,9	470	6,7	5,2	78	6,9	9,1	8,8	21,4	-10,2	8,5	9,1	+	-	HLame Only	Moderate	**RHP**	
6_C_LWP_1	6	C	LWP	1	1,5	19	1,0	1,5	21	0,3	-2,6	15,7	617	-3,7	12,1	332	4,5	-0,4	6,7	1923	3,1	2,6	85	3,1	5,3	2,4	23	-8,6	5,6	5,3	+	-	HLame Only	Mild	**RHP**	
6_A_VGP_2	6	A	VGP	2	1,5	18	0,8	1,5	19	0,3	6,5	12,6	193	1,8	10,7	597	6,8	0,4	7,0	1669	2,4	3,0	123	2,5	5,8	-4,8	15,8	-6,9	9,9	5,8	-	-			Sound	
6_B_VGP_2	6	B	VGP	2	1,5	21	0,7	1,5	21	0,3	3,2	14,6	459	4,4	9,7	218	5,5	-2,3	5,3	233	3,7	3,9	105	4,3	7,1	-6,4	14,2	-10	-4,2	7,1	-	-			Sound	
6_C_VGP_1	6	C	VGP	1	1,6	19	1,0	1,6	19	0,2	7,8	22,1	285	1,1	12,7	1164	7,8	0,8	7,3	868	0,5	2,7	569	1,0	4,9	4,4	19,4	-12,4	6,3	4,9	+	-			Sound	
8_A_LWP_1	8	A	LWP	1	1,4	23	0,9	1,4	26	0,3	-8,1	7,5	93	-2,0	14,2	712	8,4	-2,0	5,1	259	1,2	4,0	338	2,3	6,5	19,9	10,2	5,8	9,2	6,5	+	+			Sound	
8_B_LWP_1	8	B	LWP	1	1,6	19	1,3	1,6	18	0,5	-13,3	22,4	168	0,0	12,4	#####	13,3	0,3	4,4	1432	4,8	6,0	126	4,8	11,4	-7,6	16,6	-1	8,4	11,4	-	-			Sound	
8_C_LWP_2	8	C	LWP	2	1,5	17	1,1	1,5	19	0,3	-22,0	14,0	64	-0,7	21,0	2963	22,0	-2,4	4,8	198	6,2	3,3	54	6,6	17,6	4,8	8,3	-3,2	9,8	17,6	+	-	Multiple	Severe	**LFI**+RHP	2.2
8_A_VGP_1	8	A	VGP	1	1,6	17	1,1	1,6	18	0,4	-16,1	14,2	88	3,3	13,1	402	16,4	-2,5	5,4	218	2,4	4,1	170	3,5	11,7	14,4	14,2	5,8	7,3	11,7	+	+	FLame Only	Moderate	**LFI**	
8_B_VGP_1	8	B	VGP	1	1,6	18	1,1	1,6	18	0,5	-15,1	6,9	45	1,0	9,0	904	15,1	-1,9	6,2	324	4,5	6,0	135	4,9	12,4	3,7	13,4	-2,4	10	12,4	+	-	FLame Only	Moderate	**LFI**	
8_C_VGP_1	8	C	VGP	1	1,6	20	0,8	1,6	21	0,4	-8,4	12,2	145	-2,5	11,4	464	8,8	-0,6	7,8	1319	6,3	5,0	80	6,3	10,7	8,6	13,4	-5,2	10,4	10,7	+	-	HLame Only	Moderate	**RHP**	
9_A_LWP_2	9	A	LWP	2	1,5	25	0,7	1,5	25	0,3	-3,5	7,9	226	4,8	9,9	208	5,9	-4,7	6,3	134	-2,2	3,0	136	5,2	8,2	4,6	9,7	7	8,9	8,2	+	+			Sound	
9_B_LWP_2	9	B	LWP	2	1,8	18	1,2	1,8	19	0,4	-7,3	10,4	143	-2,0	14,3	732	7,5	-5,1	11,1	217	-0,2	10,8	5375	5,1	8,9	-2,1	5,8	-8,7	8,1	8,9	-	-			Sound	
9_C_LWP_2	9	C	LWP	2	1,6	20	0,8	1,6	20	0,4	-0,3	11,3	4178	2,4	9,3	390	2,4	3,5	8,2	237	5,7	3,2	55	6,7	7,9	-0,7	6,3	-5	6	7,9	-	-	HLame Only	Mild	**RHP**	
9_A_VGP_2	9	A	VGP	2	1,7	19	0,7	1,7	20	0,4	-6,2	11,1	179	2,2	5,7	258	6,6	1,3	7,0	535	6,6	5,6	84	6,7	10,0	0,9	8,2	-5	7,2	10,0	+	-	HLame Only	Moderate	**RHP**	
9_B_VGP_2	9	B	VGP	2	1,7	14	0,9	1,7	18	0,3	-10,0	7,3	73	-2,0	10,0	492	10,2	-0,7	6,0	853	1,8	4,7	256	2,0	7,1	-0,8	7,1	-3,7	7,5	7,1	-	-	FLame Only	Mild	**LFI**	
9_C_VGP_2	9	C	VGP	2	1,6	18	0,7	1,6	20	0,3	-1,7	7,8	457	4,0	13,5	334	4,4	-0,1	5,6	6233	-0,8	4,0	474	0,8	3,0	1,7	9,2	-0,1	9,5	3,0	+	-			Sound	
11_A_LWP_1	11	A	LWP	1	1,5	23	0,8	1,5	24	0,3	14,0	4,7	34	12,1	6,7	56	18,5	0,1	4,0	3636	-4,0	3,0	75	4,0	13,2	25,1	20,8	8	19,7	13,2	+	+	Multiple	Moderate	**RFI**+LHP	2.2
11_B_LWP_2	11	B	LWP	2	1,5	21	0,8	1,5	21	0,2	11,1	4,7	42	12,0	7,9	66	16,3	-2,4	3,4	142	-1,0	3,9	403	2,6	10,7	13,3	9,1	2,9	7,6	10,7	+	+	FLame Only	Mild	**RFI**	
11_A_VGP_2	11	A	VGP	2	1,6	18	0,7	1,6	19	0,4	12,1	5,9	49	-0,9	7,5	838	12,1	-3,4	10,4	305	-1,2	5,5	451	3,6	9,7	15,6	8,5	-4,1	8,2	9,7	+	-	FLame Only	Moderate	**RFI**	
11_B_VGP_2	11	B	VGP	2	1,7	18	1,2	1,7	18	0,3	27,1	8,9	33	13,0	9,7	74	30,1	-1,7	5,1	304	-2,6	4,9	187	3,1	18,2	15,2	8,6	-2	5,9	18,2	+	-	FLame Only	Severe	**RFI**	
11_C_VGP_1	11	C	VGP	1	1,6	20	1,1	1,6	20	0,4	26,2	8,2	31	11,1	5,7	52	28,5	0,9	6,2	717	-2,8	4,2	150	2,9	17,2	13,2	8,8	0,2	5,1	17,2	+	+	FLame Only	Severe	**RFI**	
13_B_LWP_1	13	B	LWP	1	1,8	10	2,8	1,8	11	0,7	-78,0	96,4	124	15,6	49,5	317	79,5	-14,2	20,1	141	-9,3	19,9	213	17,0	56,8	21,8	21	8,2	23,3	56,8	+	+			Sound	
13_B_VGP_2	13	B	VGP	2	1,6	19	1,6	1,6	19	0,3	-14,2	22,8	160	-11,8	12,2	104	18,4	-2,8	5,6	201	1,7	4,1	250	3,2	12,5	10,1	13,1	6,8	7,5	12,5	+	+			Sound	
13_C_VGP_2	13	C	VGP	2	1,6	21	2,6	1,6	23	0,3	-34,3	47,3	1380	-3,5	24,3	697	4,9	-5,6	9,0	162	-0,7	6,8	983	5,6	8,1	15,3	15,7	0	11,1	8,1	+	0			Sound	
17_A_LWP_2	17	A	LWP	2	1,5	23	2,4	1,5	25	0,4	-18,8	11,2	60	17,5	10,2	58	25,7	0,3	-10,2	6	2,8	3,2	115	2,8	15,6	-1,2	8,8	-34,2	4,9	15,6	-	-	FLame Only	Severe	LFI	
17_A_VGP_2	17	A	VGP	2	1,8	21	1,3	1,8	24	0,5	-10,8	16,9	157	10,7	11,0	103	15,2	0,5	-6,4	6	6,0	6,9	114	6,1	13,7	-1,2	8,8	-32,2	4,2	13,7	-	-			Sound	
18_B_VGP_3	18	B	VGP	3	1,7	23	1,2	1,7	22	0,3	7,3	14,7	201	-5,2	12,4	240	9,0	-5,2	5,3	8	-2,5	6,3	250	5,7	10,2	-1,2	8,8	-21,2	16,9	10,2	-	-	HLame Only	Mild	**LHI**	
22_A_LWP_1	22	A	LWP	1	1,5	22	0,9	1,5	24	0,4	-20,0	6,3	31	-2,2	7,5	335	20,1	-1,3	2,5	194	-8,6	4,5	52	8,7	18,7	-3,9	9,2	-9,7	6,1	18,7	-	-	Multiple	Moderate	**LFI**+LHP	1.2
22_B_LWP_1	22	B	LWP	1	1,6	20	0,9	1,6	22	0,4	-15,6	12,5	80	6,7	7,0	105	17,0	-2,6	4,1	156	-8,8	3,4	39	9,2	17,7	-1,7	4,1	-8,2	6,6	17,7	-	-	Multiple	Moderate	**LFI**+LHP	1.2
22_A_VGP_1	22	A	VGP	1	1,7	18	1,3	1,7	20	0,4	-10,5	15,0	142	0,3	11,5	4600	10,5	-4,7	5,5	117	-8,4	5,1	60	9,6	14,9	-0,8	10,5	-11,7	9,6	14,9	-	-	HLame Only	Moderate	**LHP**	
22_B_VGP_1	22	B	VGP	1	1,7	18	1,2	1,7	18	0,6	-14,1	21,4	152	-15,0	25,9	173	20,6	-6,7	7,7	115	-14,0	5,8	41	15,5	25,8	6,6	20,9	-13,3	10,4	25,8	+	-	HLame Only	Severe	**LHP**	
23_A_LWP_2	23	A	LWP	1	1,7	23	1,9	1,7	23	0,4	-48,9	32,2	66	-4,2	14,8	349	49,0	3,8	7,3	192	-4,1	4,2	102	5,6	30,1	-0,4	6,8	0,9	3	30,1	-	+	FLame Only	Severe	**LFI**	
23_B_LWP_1	23	B	LWP	1	1,6	25	1,7	1,6	24	0,3	-38,4	26,4	69	-0,4	14,9	3911	38,4	4,1	6,3	153	-2,2	3,4	155	4,7	23,8	-6,7	6,5	2	4,3	23,8	-	+	FLame Only	Severe	**LFI**	
23_C_LWP_2	23	C	LWP	2	1,7	24	1,7	1,7	24	0,3	-49,2	22,5	46	6,5	20,4	313	49,6	4,4	7,1	162	-1,3	6,2	463	4,6	29,3	-1,3	8,1	1	4,9	29,3	-	+	FLame Only	Severe	**LFI**	
23_A_VGP_1	23	A	VGP	1	1,7	18	1,8	1,7	18	0,3	-45,4	23,5	52	3,3	11,9	356	45,5	3,3	7,5	230	-0,9	7,0	803	3,4	26,1	-7,5	10,9	4,8	4,5	26,1	-	+	FLame Only	Severe	**LFI**	
23_B_VGP_1	23	B	VGP	1	1,8	18	1,4	1,8	19	0,4	-31,2	15,6	50	5,1	10,2	198	31,7	5,0	8,0	160	-3,1	40,9	1311	5,9	21,7	-2,9	7	4,5	8,5	21,7	-	+	FLame Only	Severe	**LFI**	
23_C_VGP_1	23	C	VGP	1	1,7	19	2,1	1,7	20	0,4	-33,0	32,1	97	14,1	15,2	108	35,9	4,0	8,3	208	-6,2	8,7	141	7,3	25,3	-3,2	10,4	5,6	7	25,3	-	+	FLame Only	Severe	**LFI**	
28_C_LWP_2	28	C	LWP	2	1,7	18	1,5	1,7	20	0,3	25,9	11,7	45	22,9	7,5	33	34,6	3,1	4,6	150	-3,1	6,4	209	4,3	21,6	4,9	6,3	1,7	3,9	21,6	+	+	FLame Only	Severe	**RFI**	
28_B_LWP_2	28	B	VGP	2	1,7	21	1,4	1,7	21	0,3	18,5	8,4	45	12,7	9,0	71	22,5	5,6	3,0	53	-4,5	5,0	111	7,2	18,4	5,3	5,7	-1,2	3,7	18,4	+	-	Multiple	Severe	RFI+**RHI**	1.3
28_C_VGP_2	28	C	VGP	2	1,8	17	1,6	1,8	18	0,3	32,4	14,4	44	19,3	8,0	41	37,7	7,3	5,5	75	-3,2	5,2	164	8,0	26,8	5,8	5,4	-0,3	5,6	26,8	+	-	Multiple	Severe	RFI+**RHI**	1.3
29_A_LWP_1	29	A	LWP	1	1,5	23	1,1	1,5	23	0,4	21,1	9,6	45	4,9	4,3	87	21,7	-10,4	4,7	45	-5,2	6,9	132	11,6	22,4	16,4	6,2	13,2	3,7	22,4	+	+	Multiple	Severe	**RFI**+LHI	2.3
29_B_LWP_1	29	B	LWP	1	1,5	21	0,7	1,5	21	0,42	12,7	12,8	101	2,4	4,8	199	12,9	-9,9	2,4	24	-7,5	4,4	60	12,4	18,9	13,8	6,1	11,1	2,6	18,9	+	+	Multiple	Moderate	**RFI**+LHI+P	2.1
29_C_LWP_1	29	C	LWP	1	1,5	21	0,8	1,5	21	0,51	18,2	12,7	70	4,5	6,9	155	18,7	-12,1	4,0	33	-8,9	4,3	48	15,0	24,3	17,4	7,8	12,7	2,2	24,3	+	+	Multiple	Severe	**RFI**+LHI+P	2.1
29_A_VGP_1	29	A	VGP	1	1,6	21	1,6	1,6	21	0,56	-8,2	27,3	335	-7,0	18,8	268	10,8	-13,5	8,1	60	-5,9	5,8	99	14,7	20,1	9,1	7,1	11,4	7,4	20,1	+	+	HLame Only	Severe	**LHI**	
29_B_VGP_2	29	B	VGP	2	1,6	19	1,1	1,6	19	0,46	19,3	14,9	77	3,0	8,8	296	19,5	-8,7	4,9	56	-7,4	7,1	95	11,4	21,2	15,1	7,2	10,9	6,8	21,2	+	+	Multiple	Severe	**RFI**+LHI+P	2.1
29_C_VGP_2	29	C	VGP	2	1,6	20	0,9	1,6	20	0,55	13,3	12,1	91	1,1	12,5	1123	13,3	-14,5	6,4	44	-5,4	5,1	94	15,5	22,1	12,2	7,3	15,1	5,8	22,1	+	+	Multiple	Moderate	**RFI**+LHI+P	2.1
30_A_LWP_1	30	A	LWP	1	1,5	14	1,1	1,5	17	0,34	20,6	10,3	50	10,2	9,7	95	22,9	-2,5	4,8	188	-5,8	4,3	75	6,3	17,8	6,7	4,1	4,3	9,1	17,8	+	+	Multiple	Severe	**RFI**+LHI+P	2.1
30_B_VGP_2	30	B	LWP	2	1,5	19	0,9	1,5	20	0,22	19,2	25,4	133	2,6	12,6	484	19,4	-1,1	3,8	359	-3,6	3,2	87	3,8	13,5	5,0	4,3	-1	5,5	13,5	+	-	HLame Only	Mild	**LHP**	
30_C_LWP_2	30	C	LWP	2	1,6	15	1,9	1,6	22	0,28	6,8	22,9	339	8,3	7,6	91	10,7	-1,8	6,6	368	-3,8	3,1	81	4,2	9,6	10,6	4,8	0,2	3,5	9,6	+	+	HLame Only	Mild	**LHP**	
30_A_VGP_3	30	A	VGP	3	1,5	14	2,4	1,5	19	0,53	42,4	29,5	70	15,8	23,0	146	45,3	0,1	16,8	12908	-1,7	12,7	749	1,7	24,3	6,7	11,3	3,7	9,5	24,3	+	+	FLame Only	Severe	**RFI**	
30_C_VGP_2	30	C	VGP	2	1,6	19	1,4	1,6	19	0,36	13,2	26,2	199	-2,3	22,0	974	13,4	-3,4	4,1	121	-6,7	4,2	63	7,5	14,2	8,1	3,9	2,5	8,3	14,2	+	+	HLame Only	Moderate	**LHP**	
31_B_LWP_2	31	B	LWP	2	1,6	20	0,5	1,6	20	0,31	4,1	8,6	212	1,6	10,4	659	4,4	2,3	3,8	165	6,9	4,6	67	7,2	9,4	1,4	6,3	9,3	3,2	9,4	+	+	HLame Only	Moderate	**RHP**	
31_C_LWP_2	31	C	LWP	2	1,7	19	0,8	1,7	19	0,43	13,1	10,2	77	0,1	8,5	8490	13,1	2,8	4,9	173	10,6	4,5	42	10,9	17,5	3,4	4,7	4,2	4,4	17,5	+	+	Multiple	Moderate	**RFI**+RHP	1.2
31_A_VGP_1	31	A	VGP	1	1,8	23	1,3	1,8	23	0,53	-7,6	23,8	315	-13,4	17,8	133	15,3	0,1	7,4	6745	2,1	11,6	567	2,1	9,7	7,4	9,5	-2	10	9,7	+	-			Sound	
31_B_VGP_2	31	B	VGP	2	1,7	20	1,1	1,7	22	0,42	11,5	14,1	122	-1,1	10,8	967	11,6	-0,8	5,8	729	9,5	7,0	73	9,5	15,3	7	3,9	6,5	5,8	15,3	+	+	HLame Only	Moderate	**RHP**	1.2
31_C_VGP_1	31	C	VGP	1	1,7	20	0,7	1,7	20	0,37	0,3	11,5	3371	0,3	8,8	2917	0,5	1,8	4,2	231	3,5	8,3	236	3,9	4,2	8,4	7,7	7,6	4,1	4,2	+	+			Sound	
32_A_LWP_1	32	A	LWP	1	1,6	15	2,0	1,6	20	0,47	-5,2	20,4	391	-0,9	8,9	955	5,3	-2,2	3,8	172	8,6	8,7	101	8,9	11,6	-1,2	9,4	-1,8	5,6	11,6	-	-			Sound	
32_A_LWP_2	32	A	LWP	2	1,7	17	1,5	1,7	17	0,42	-4,4	23,3	530	-3,1	24,2	794	5,3	-0,7	5,9	903	7,3	7,9	108	7,3	10,0	0,6	8,7	5,5	13,2	10,0	+	+			Sound	
32_B_LWP_1	32	B	LWP	1	1,6	16	2,4	1,6	19	0,35	13,3	42,2	316	-4,8	41,1	849	14,2	0,8	5,6	663	8,8	4,3	49	8,8	15,9	4,6	7,9	3,1	11,5	15,9	+	+	HLame Only	Moderate	**RHP**	
32_C_LWP_2	32	C	LWP	2	1,6	17	1,8	1,6	17	0,39	26,5	34,3	130	-6,6	18,0	274	27,3	-4,6	7,1	156	9,2	4,9	53	10,3	23,9	9,1	9,1	-1,2	10,2	23,9	+	-	HLame Only	Moderate	**RHP**	
32_A_VGP_2	32	A	VGP	2	1,7	16	2,6	1,7	19	1,26	19,8	36,6	185	22,0	20,6	94	29,6	-9,2	65,4	713	8,9	53,6	605	12,8	27,5	-6,0	47	14,7	15,7	27,5	-	+	FLame Only		**RFP**	
32_B_VGP_1	32	B	VGP	1	1,62	15	1,7	1,6	16	0,4	12,31	43,0	349	-7,1	27,7	389	14,2	-4,66	9,8	209,9	5,95	4,86	82	7,56	14,67	7,9	5,4	5,2	8,9	14,67	+	+	HLame Only	Mild	RHP	

CHAPTER VI: Salivary Cortisol and Infrared Thermographic Ocular Temperature Use as Biomarkers During Endurance Competitions

Mónica C. de Mira, Elsa Lamy, Rute Santos, Jane Williams, Mafalda Vaz Pinto, Pedro Martins, Patrícia Rodrigues, David Marlin (2021). Submitted to BMC Veterinary Research (second revision). Pre-print available at: https://www.researchsquare.com/article/rs-111435/v1

de Mira *et al. BMC Veterinary Research* (2021) 17:329
https://doi.org/10.1186/s12917-021-02985-9

BMC Veterinary Research

RESEARCH ARTICLE Open Access

Salivary cortisol and eye temperature changes during endurance competitions

Monica C. de Mira[1*], Elsa Lamy[1], Rute Santos[1,2], Jane Williams[3], Mafalda Vaz Pinto[1], Pedro S. Martins[4], Patrícia Rodrigues[5] and David Marlin[6]

Abstract

Background: The purpose of this study was to investigate the usefulness of salivary cortisol (SC) and eye temperature measured by infrared thermography (IRT^{ET}) as biomarkers to manage competitions more effectively and monitor horse welfare in endurance competitions. Based on previous studies, it was hypothesised that pre-exercise baseline SC and IRT^{ET} would be higher in younger or less experienced horses, and that post-exercise variation from baseline would be higher in the top finishers.

Results: Salivary cortisol measured in 61 competing at qualifier 40 km and 80 km rides showed an abrupt variation (93–256% rise) of the baseline SC levels [median ± interquartile range (IQR) = 0.27 ng/dl ± 0.36] obtained at the Pre-Inspection (PI) into Vet Gate (VG)1 independently of the covered distance, but modest or even lower in the subsequent Vet Gates, e.g. VG2 or VG3. The IRT^{ET} measured concomitantly in 16 horses showed significant ($p < 0.05$) higher levels at the PI in less experienced horses participating in the 40 km ride (median ± IQR = 35.7 °C ± 1.4) than their counterparts in the 80 km ride (median ± IQR = 35.0 °C ± 1.5), but not SC. Baseline SC levels at the PI of horses classifying in the Top5 in the 40 km ride category were significantly ($p < 0.05$) higher median ± IQR = 0.90 ng/ml ± 0.61) when compared to horses positioned from 10th position on (median ± IQR = 0.16 ng/ml ±0.40). A lower IRT^{ET} in the PI was correlated with better placement ($p < 0.05$) and those in the Top5 (median ± IQR = 33.9 °C ± 0.0) had a significantly ($p < 0.5$) higher variation (+ 10.65%) into the last VG.

Conclusion: Pre-exercise baseline IRT^{ET} levels, but not SC, were higher in less experienced horses in the 40 compared to their counterparts in the 80 km ride competitions. SC and IRT^{ET} showed different indications according to the competition. In the40 km ride competition, higher baseline pre-exercise SC levels seemed to be linked to a better classification outcome. In contrast, in the 80 km ride horses, the higher IRT^{ET} variation from pre-exercise into final Vet Gate was the parameter associated with a better performance. A more controlled environment and a larger sample are needed to confirm these results and monitor horse welfare in competitions.

Keywords: Endurance riding, Eye temperature, Infrared thermography, Salivary cortisol, Performance, Equine

* Correspondence: monicademira@gmail.com
[1]MED – Mediterranean Institute for Agriculture, Environment and Development, Institute for Advanced Studies and Research, Universidade de Évora, Pólo da Mitra, Ap. 94, 7006-554 Évora, Portugal
Full list of author information is available at the end of the article

Background

Endurance ride competitions are long-distance races of 40 to 160 km against the clock in phases that consist of a minimum of 16 to a maximum of 40 km, followed by a required rest period, at least equal in minutes to the distance in km of the competition [1]. Mandatory veterinary inspections, before (pre-inspection) and after each phase are performed is an assigned area called the Vet Gate (VG) to determine if competing horses are fit to compete or need to be eliminated to protect their integrity [2]. Despite the highest elimination rates among all equestrian sports [3–5] and the introduction of stricter rules by the Fédération Equestre Internationale (FEI) severe injuries still occur. This is not only unacceptable for today's societal welfare standards towards equine athletes [6, 7] but also frustrating for veterinarians, who are often confronted by competitors with the subjectivity of a decision to eliminate a horse [8]. For these reasons and because horses, unlike humans, cannot vocalise distress or pain or make decisions for themselves, the possibility of utilising non-invasive and objective methods, such as gait sensors [9] and biomarkers would be instrumental to evidence-based management of the equine athlete's welfare while competing. Cortisol determination using saliva and eye temperature measured by infrared thermography are non-invasive techniques used to evaluate horses' stress responses to its human equestrian utilisation [10].

Exercise is a naturally a stressor per se and induces a biologic response that can be either an enhancer or a limiting factor for an athlete's sporting ability [11]. Yet, in competitions, equine athletes face a mixture of other stressors including transportation [12], veterinary examinations [13], rider's ability [14], a new and a noisy environment [15], separation from stablemates [16] and, specifically in endurance, exposure to large conglomerations of unfamiliar horses in large starts, and musculoskeletal pain from an injury that might arise [17] that can also elicit a stress response. Moreover, individual intrinsic factors such as age, gender, breed, inherited temperament, experience, and previous training [11] are known to impact stress biomarkers.

Cortisol production is the end-result of the hypothalamic-pituitary-adrenal (HPA) axis activation induced by any psychological or physical stressor. It has been studied extensively in horses to quantify stress levels and the response to different types, intensities and durations of exercise in sport and racehorses [10, 18]. The validation of salivary cortisol (SC) [19] allowed its non-invasive assessment in competition settings, including endurance [20–22], showjumping [15, 23, 24] and dressage. Even if a circadian rhythm has been demonstrated, plasma levels did not always correlate with salivary cortisol [25]. The difference could be explained by only the biological active unbound component being present in saliva, whereas, in plasma, both the inactive and active free constituents of cortisol are measured and not necessarily proportional [26]. Cortisol showed greater variations in saliva than in plasma [19, 27]. The highest variations from pre-exercise in salivary cortisol (SC) levels were registered in endurance (up to 1000%) [20] followed by eventing (240%) [28], showjumping (150–340%) and dressage (200%) [29] competitions.

The changes in circulation associated with the HPA axis activation induce periorbital warming that can be quantified by thermal imaging cameras [10]. The use of hairless vascularised areas such as the lacrimal caruncle to measure temperature minimises interference from skin and coat colour, and environmental conditions [30]. The rise in eye temperature measured by infrared thermography (IRT^{ET}) was reported as a reliable indicator of short-term stress, and is often studied together with salivary cortisol in horses [10, 14, 31–34]. It has been generally accepted that a rise in eye temperature represents an emotional response to stressors, including exercise [35], as opposed to a physiological response to exercise's physical demands, as proposed recently [31]. IRT^{ET} may represent a measure of emotive reactivity to effort, that can have a beneficial or detrimental effect on performance [14, 35]. For this reason, IRT^{ET} has recently been proposed as a selection tool to help identify emotional reactivity as a desirable, or undesirable, trait for performance according to the horse's intended use [35, 36]. The complimentary use of salivary cortisol and IRT^{ET} as non-invasive biomarkers of stress during endurance competitions could help characterise distress and physiological response to effort for endurance horses during exercise in competition.

To our knowledge IRT^{ET} alone or concomitantly with SC has not been studied before during endurance rides. This study aimed to investigate the usefulness of these biomarkers to manage competitions more effectively and monitor horse welfare in endurance competitions. Based on previous studies, it was hypothesised that pre-exercise baseline SC and IRT^{ET} would be higher in younger or less experienced horses, and that post-exercise variation from baseline would be higher in the top finishers.

Results

Horse's previous experience and competition outcome

Age was not significantly different between the 40 km (40K) (median ± IQR = 6.0 ± 1.5) and 80 km (80K) categories (median ± IQR = 6.0 ± 3.0), however, there was a significant ($p < 0.05$) difference in previous experience between the two categories: horses in the 40K had less km in competitions (median ± IQR = 40 ± 30, min = 0, max 120) than horses in the 80K category (median ± IQR =

80 ± 40, min = 80, max = 240). Across all competitions, a total of 11 horses (18%) failed to qualify, six for irregular gait, two for metabolic reasons and the remaining three for other reasons. The speed median (±IQR) was 14.9 km/h (±2.5) and 15.7 km/h (±1.0) for the 40 and 80K categories, respectively. In the first phase, horses in the 40K covered 20 km at a significantly ($p = 0.006$) slower speed (median ± IQR =14.0 km/h ± 1.8), when compared with those in the 80K ride category, that covered either 30 or 40 km (median ± IQR =15.1 km/h ± 0.9).

Visual assessment of saliva samples

The saliva samples were subjectively judged to have less volume with the progression of the ride. Also, many samples were contaminated with food particles that horses kept in the mouth during the ride phases.

Age and gender impact in SC and IRT^{ET}

No significant differences or correlations were identified between SC or IRT^{ET} with age or gender, except for mares that showed a significantly higher SC ($p = 0.037$) at the final VG3 after 80 km covered (VG3@80km) in the 80 K-B ride.

SC and IRT^{ET} measurements

Means and medians of SC and IRT^{ET} of all individuals collected at the different moments (previous and competition day) are displayed in Fig. 1.

Baseline values

The lowest SC levels were registered in all categories at Home or PI. When comparing the 40 K with the 80 K horses' baseline SC values, there was not a significant difference at Home nor the PI. At the PI, IRT^{ET} was, however, significantly higher ($p = 0.007$) in the horses competing in the 40K (median ± IQR = 35.7 ± 1.4) when compared to those in the 80K (median ± IQR = 35.0 ± 1.5) category ride (Table 1).

Analysis by vet gate

The highest SC levels were registered in the first Vet Gate after 30 or 40 km covered (VG1@30/40km) in the

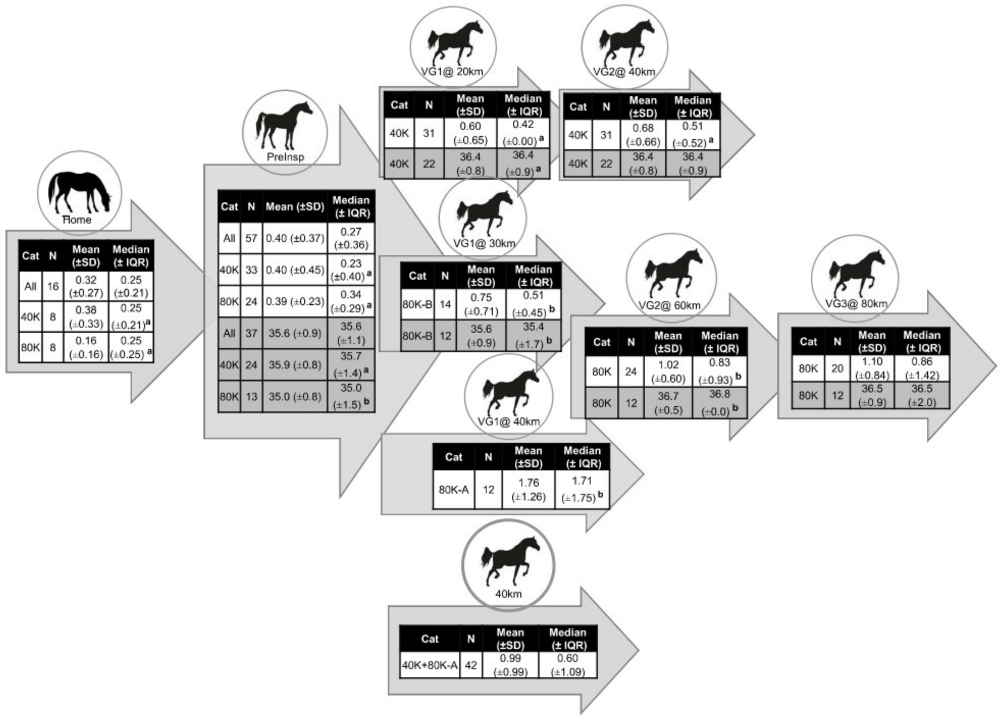


Fig. 1 Means and medians of salivary cortisol in ng/dl (white background) and Infrared Thermographic Eye Temperature in Celsius degrees (grey background) of all horses (ALL), in 40K and 80K categories, collected at Home, Preinspection (PreInsp), by Vet Gate (VG) and covered distance (@). Significant differences ($p < 0,05$) among ride categories at Vet Gates (columns) are signalised by different letters

Table 1 Correlations of Salivary Cortisol (SC) and OT Eye Temperature (ET) measured by Infrared Thermography with performance data in the 40 K and 80 K ride categories

Spearman Rank Correlation				
SPEED	**Rho**	**N**	***P*-Value**	**Ride Cat**
PHASE 1				
ΔSC Home-PI	- 0,7*	10	0,023	40 km
ET PI	- 0,6	13	0,024	80 km
ΔET PI-VG60 km	+ 0,6*	12	0,027	80 km
ΔET VG60- VG80 km	+ 0,6*	12	0,037	80 km
PHASE 2				
ΔSC PI-VG30 km	+ 0,6	13	0,047	80 km
ET VG3	+ 0,6	12	0,045	80 km
ΔET PI -VG60 km	+ 0,7*	12	0,015	80 km
ΔET VG30 -VG 60 km	+ 0,7*	10	0,033	80 km
PHASE 3				
ET VG3	+ 0,6*	12	0,12	80 km
ΔET PI-VG60 km	+ 0,7**	12	0,009	80 km
ΔET PI-VG80 km	+ 0,7*	12	0,033	80 km
AVERAGE				
ET PI	- 0,6*	11	0,038	80 km
ET 20 km	- 0,5*	22	0,024	40 km
ET 80 km	+ 0,7*	11	0,024	80 km
ΔET PI-VG60 km	+ 0,8**	11	0,006	80 km
ΔET PI-VG80 km	+ 0,7*	11	0,035	80 km
RECOVERY TIME				
VG1				
ET 60 km	- 0,6*	12	0,034	80 km
ΔET VG0-VG80 km	+ 0,6*	12	0,036	80 km
VG3				
ΔSC VG0-VG30 km	- 0,6*	13	0,047	80 km
ΔET VG30-VG80 km	- 0,8**	10	0,009	80 km
ΔET VG60-VG80 km	- 0,6*	12	0,028	80 km
QUALIFICATION				
CL vs FTQ				
ET PI	−0,5*	24	0,026	40 km
POSITION				
SC PI	−0,5*	25**	0,025	40 km
ET VG80 km	−0,7*	12	0,011	80 km
ΔET PI-VG60 km	−0,8**	12	0,005	80 km
ΔET PI-VG80 km	−0,6*	12	0,034	80 km

p-value is two-tailed * ($p < 0{,}05$) **($p < 0{,}001$). PI (Pre-Inspection), VG (Vet Gate); 20, 30, 60 and 80 (distance covered in km. Δ (variation between moments of collection)

80K ride category, but only in the second or final Vet Gate after twice twenty km covered (VG2@40km) in the 40 K category. In contrast, the highest IRT^{ET} was obtained in VG2 in both ride categories, independently of the covered distance, i.e. 40 or 60 Km. When comparing ride categories at VG1, the 40K horses had a significantly lower SC ($p = 0.006$), but a significantly higher IRT^{ET} ($p = 0.023$), than the 80K horses.

Analysis by covered distance

When comparing the same covered distance among ride categories, horses in the 40K having performed two phases of 20 km, with a rest period in-between, showed in VG2 a significantly ($p = 0.001$) lower SC, when compared with those in the 80K ride that had raced uninterruptedly 40 km in one phase and were at VG1.

Analysis by final outcome (completion vs failing to qualify: FTQ)

There were no significant differences in SC or IRT^{ET} measurements between the horses that completed the ride and those that failed to qualify, in none of the evaluated moments.

Analysis by classification group

SC or IRT^{ET} levels analysed by classification groups and its evolution across Vet Gates in both ride categories, e.g. horses positioned in the top five (Top5), from 6th to 10th (G2) and from 11th to 15th (G3) can be consulted in Fig. 2. Horses in the 40K competition classifying in the Top5 showed at the PI, significantly higher (p = 0.05) SC levels (median ± IQR = 0.90 ng/ml ±0.61) when compared to horses positioned in G3 (median ± IQR = 0.16 ng/ml ±0.40). On the other hand, horses classifying in the Top5 in the 80K competition, demonstrated at VG2, significantly ($p = 0.05$) lower SC levels (median ± IQR = 0.70 ng/ml ±1.00) than horses positioned in G2 (median ± IQR = 1.88 ng/ml ±1.00) and at VG3 (final), a significantly (p = 0.053) higher IRT^{ET} (median ± IQR = 37.60 °C ± 0.00), than horses positioned in G3 (median ± IQR = 35.70 °C ± 1.00).

Variations of SC and IRT^{ET} between collection moments

The magnitude of variations in SC levels and IRT^{ET} between Home, PI and Vet Gates of different ride categories and its significance is shown in Fig. 2. The baseline SC variation between Home and next day PI was only significant ($p = 0.017$) for those horses participating in the 80 K ride category with a 122% rise. The highest SC variation was between PI and VG1, but only significant, in the 80K ride category with a 216 and 256% rise in 80 K-A and 80 K-B, respectively. IRT^{ET} rises were only significant when values were compared across more than one Vet Gate. When analysed by classification group, horses classified in the Top5 (median ± IQR = 33.9 ± 0.0) and in G3 (median ± IQR = 35.3 ± 1.0) had a variation of 10.65 and 1.78% from the PI to VG3, respectively (Fig 3).

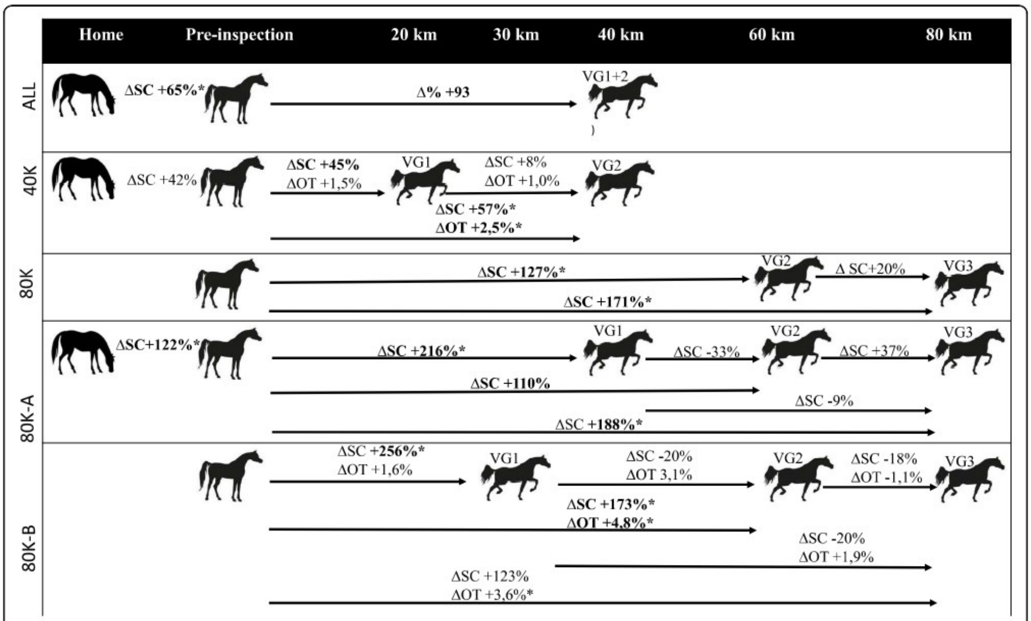


Fig. 2 Variations of SC (Salivary Cortisol) and IRTET (Eye Temperature measured by Infrared Thermography) of all horses (ALL), 40 km and 80 km categories, from and in-between Home, Preinspection and covered distances at each Vet Gate. The variations represented between Home and Preinspection represent the sample of horses collected both at home and at the pre-inspection. The variations between Pre-Inspection and covered Km represent the same sample of horses collected at the respective covered distances, thus eliminated horses or horses that failed a collection were not compared. Significant variations among coincident points-in-time and/or covered distance of all horses (ALL), 40 km, 80 km, 80 km-A and 80 km-B categories are in bold signalised by * ($p < 0{,}05$)

Correlations between SC, IRT^{ET} and performance

Significant correlations between SC or IRT^{ET}, its variations and performance data (speed, recovery time, final position) are presented in Table 1. An association could not be established between SC baseline values, nor its variations from Home to PI, with the outcome. However, when analysed by classification group, a classification in the Top5 of the 40K ride category was significantly ($p < 0.05$) associated with an IRT^{ET} decrease from PI to VG2. In contrast, in the 80K ride category, a lower IRT^{ET} at the PI was significantly ($p < 0.05$) associated with a faster speed in phase 1, overall average speed and completion. Also, the higher the IRT^{ET} variation from PI into VG2 and VG3 was associated with a better placement.

Correlations between SC and IRT^{ET} were scarce and are depicted in Table 2.

Discussion

Endurance riding evolved in the last two decades from an amateur activity into a highly professionalised sport. Better training techniques and more specialised breeding allowed the creation of equine endurance super-athletes, capable of achieving a sustained high speed along with a fast-cardiac recovery capacity. This preliminary study aimed to determine how salivary cortisol (SC) and eye temperature measured by infrared thermography (IRT^{ET}) and their variations before and during endurance competitions were related to the outcome and performance of competing horses, and their potential usefulness in depicting compromised horses.

Behaviour of SC and IRT^{ET} during competitions

Various factors inherent to competitions, such as accustoming to a novel environment [37] and a new group of horses [38] or undergoing a veterinary examination [39] have been described as potential stressors to horses. Transportation is considered a major stressor capable of generating greater SC rises than exercise [13]. Even in short distances such as 1 h, a 4-fold SC increase was previously reported [12]. All horses in our study were transported to the venue the same morning of the competition, arriving typically near the time of the PI and the estimated transportation time ranged between 10 min and no longer than 2 h. In this study, an overall 65% SC increase from the baseline values the eve at Home to the

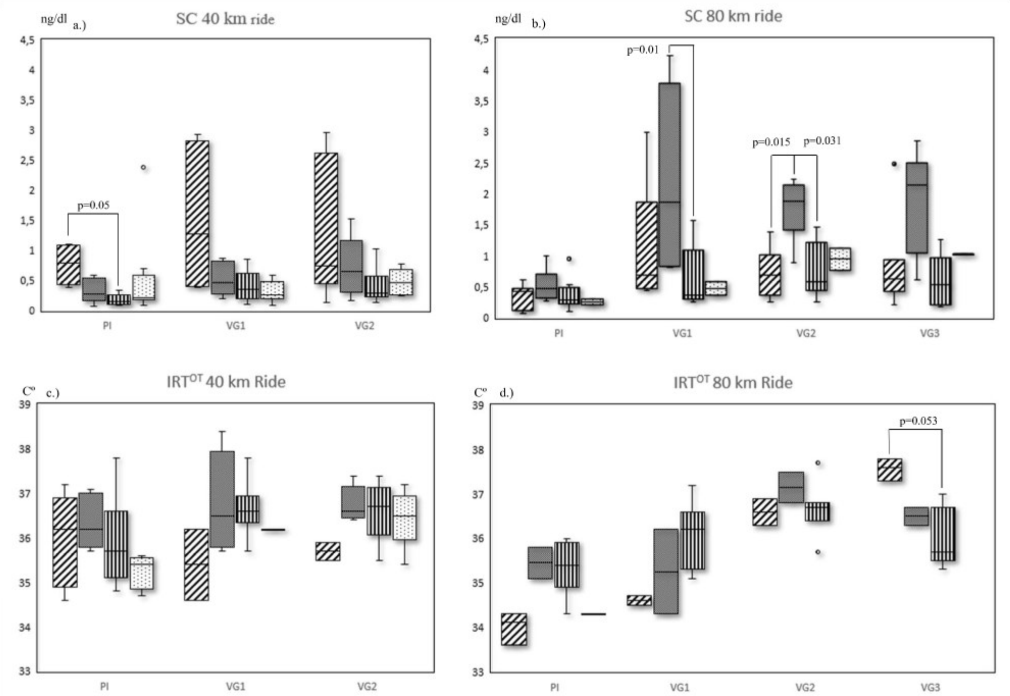


Fig. 3 Clustered boxplots of SC (Salivary Cortisol) and IRT^{ET} (Eye Temperature measured by Infrared Thermography) by group position. a.) Salivary Cortisol (SC) in 40 and b.) 80 km rides, and c.) Infra-Red Thermography Eye Temperature (IRT^{ET}) in 40 and b.) 80 km rides with horses grouped in Top5 (1th–5th), G2 (6-10th), >11th and FTQ (Failed to Qualify), aligned from left to right, at Preinspection (PI) and Vet Gates (VG).Horizontal X-Axis: Vet Gates (VG); Vertical Y-Axis: SC in ng/dl and IRT^{ET} in centigrade C°

Table 2 Significant correlations found between SC (Salivary Cortisol) and Eye Temperature (ET) measured by Infrared Thermography in the 80 km category

Spearman Rank Correlation			
	Rho	**N**	***P*-Value**
SC 40 km			
ET 60 km	- 0,9*	5	0,037
Δ SC VG30-VG60 km			
ΔET PI-VG30 km	- 0,8*	8	0,032
ΔETET PI -VG80 km	- 0,8*	9	0,012
ΔSC VG40-VG80 km			
ΔET VG60-VG80 km	+ 0,8*	6	0,050
ΔSC VG60-VG80 km			
ΔET VG60-VG80 km	- 0,8*	6	0,050

p-value is two-tailed * ($p < 0,05$) **($p < 0,001$) PI-Pre-Inspection), VG (Vet Gate); 20, 30, 60 and 80 km (distance covered in km. Δ (variation between moments of collection)

first collection performed at the competition venue immediately after the PI, was modest and less than the competition itself's rise.

Higher cortisol rest levels [13, 24] and IRT^{ET} [33, 35] were previously reported in younger or less experienced horses. However, basal cortisol levels have also been reported to be similar in a competition setting between horses with different experience levels [40]. Even if there was not a significant difference in SC levels at Home or the PI among ride categories, in our case, IRT^{ET} was higher in the less experienced horses participating in the 40 K ride in the PI. Both SC and IRT^{ET} have been used as indicators of distress in non-exercised horses [34]. Eye temperature is considered a more immediate stress indicator than cortisol, reported to take at least 15 min to increase after exposure to a stressor [10]. Since IRT^{ET} was measured immediately after exiting the VG, this could reflect a higher distress of the 40 K horses exposed to the veterinary examination and, often, being separated from their mates at the PI.

As expected, both SC and IRT^{ET} lowest values were registered at Home and at the PI. However, regarding the occurrence of the highest values there was a difference between ride categories. In the 40K ride the maximum SC and IRT^{ET} were registered in the final Vet Gate, as opposed to the 80K rides, where they were obtained at mid-distance in VG1 and VG2, respectively but not in the final Vet Gate.

Our study corroborates that both intensity and duration, if uninterrupted, contribute to SC increase [41]. Indeed, those horses in the 80 K category that performed a straight 40 km phase into VG1 at a higher speed showed a 3-fold higher cortisol level than horses in the 40K ride, which raced two 20 km phases with a rest period in-between.

Our study agrees with previous studies performed during endurance competitions, which also registered the highest SC increases in the first half of the rides [20–22]. Regardless of the covered distance and registered levels, the steepest SC variations took place in both ride categories in VG1, showing much more modest, or even negative variations, in the subsequent VGs. This effect was also reported previously in human athletes, whose cortisol levels increased after short-term and decreased after prolonged, i.e. lasting several hours, exercise [42]. This drop is believed to result from the negative feedback system generated by the high cortisol levels induced by exercise. Two mechanisms were proposed for athletes to bypass the negative feedback. First, interleukin-6 released from working muscles induced by low glycogen contents seems to act as a hormone, stimulating, similarly to cortisol, the maintenance of glucose homeostasis during exercise and mediating exercise-induced lipolysis [43]. The second mechanism could be the individual's inherent ability to override the serotonergic mechanisms (that inhibit CRH release and therefore the HPA axis) involved in central fatigue, which is not necessarily related with training level [44]. This drop could also be connected to a decrease in the first moment from a decrease in horses' emotional stress content. It was also proposed before that the initial higher levels could be associated with excitement and not with body demand [20]. The emotional stress could, therefore, explain the variations reported in other studies at similar magnitudes, but in much lighter exercises [29].

IRT^{ET} was used to characterise stress levels induced by certain equestrian practices such as neck hyperflexion [14] or a tight noseband [45]. More recently, IRT^{ET} was also studied in showjumping [33, 46] and dressage competitions [36], in Standardbred harness races [35] and in flat race Arabian and Thoroughbred horses in training [32]. One of the proposed added values of the use of IRT^{ET} is its potential independence from the effort effect, thereby providing a valid means of evaluating the emotive reaction to effort stressors in exercised horses [10].

We could only find very few associations between SC and IRT^{ET}. This is in line with other studies that investigated SC and IRT^{ET} simultaneously during exercise [14, 32, 33, 35, 47]. One study could establish an association between the two biomarkers in exercise, but only after an ACTH stimulating challenge test [48] and another, during clipping, a non-exercise activity [34].

In our study, the highest IRT^{ET} rise in consecutive Vet Gates was from VG1 to VG2 (+ 3.1%) in the 80 K ride, that also corresponded to the highest SC drop (− 20%). This might be explained by cortisol representing mainly the physiological response to exercise, and the eye temperature, the prolonged effort's emotional reactivity.

SC levels and IRT^{ET} association with competition outcome

Elevations of basal cortisol concentrations in response to emotional stress are believed to be detrimental to general health, but not necessarily to sport performance [49]. Indeed, in the more inexperienced horses of the 40K ride, the higher SC levels before and during the ride were associated with better performance, reflecting most likely the extra necessary physiological response to effort (Table 1 and Fig. 1). In the 80 K category, cortisol behaved differently. It appears it was not the pre-exercise SC level that influenced the results per se, but the magnitude of increase from PI to VG1@30km associated with a higher placement group (Table 1). Moreover, the group finishing in the Top5 showed a significantly lower SC than the slower G2 in the second-to-last vet gate or VG2. This may indicate an extra effort in less well-prepared horses of G2. Cortisol was shown to increase with effort intensity, but in horses subjected to the same amount of exercise, the rise was higher in untrained horses [50].

IRT^{ET} was proposed as an alternative biomarker capable of quantifying emotional reactivity to effort, instead of a direct measure of effort like cortisol [35, 36, 46]. A lower and higher IRT^{ET} before and after exercise, respectively, i.e., a higher variation after exercise, was reported to be associated with better performances by analysing 130 Spanish Standardbred horses in harness races [35]. The same authors concluded that a variation of − 0.97% represented the break-point under which physiological stress developed. In our study, the 80K category horses with a lower IRT^{ET} at the PI and a more significant rise into the final VG3 were better placed in the final classification (Table 1). Furthermore, this rise was associated with a shorter recovery time in VG3, but not in VG1, which might be attributed to the initial excitement. The 40K ride horses showed very few associations with IRT^{ET}. A reason for that could be that they started with an already higher IRT^{ET} at the PI. Negro et al. [35] estimated a pre-race eye temperature of

37.61 ± 2.85 C °C with a post-race variation of + 7.57% as the optimal values for performance. Our lower number of horses and the prolonged low-intensity non-explosive nature of endurance exercise, precluded these calculations. Yet, horses classified in the Top5 when compared with G3 had an average IRT^{ET} of 33.8 °C and 35.33 °C with a variation of + 10.65% and + 1.78%, respectively.

More studies are warranted to investigate the meaning and usefulness of IRT^{ET}.·A recent study proposed IRT^{ET} as an indicator of physical fitness in ranch horses [31], as opposed to a purely psychological reaction to effort. The rise was attributed to increased blood flow in muscles and peripheral heat dissipation. A correlation was found with creatine kinase (CK), indicating a possible association with muscle damage.

Failure to qualify

In this research, most likely due to the small sample, we could not find a difference or association between eliminated or classified horses and SC levels or IRT^{ET}.

Limitations of the study

Volume and food contamination in SC determination In this study, we used the saliva collection protocol described by Peeters et al. (2001). Therefore, in further endurance studies, we recommend that due to the horses' progressive natural dehydration, which likely justified the diminished saliva volume observed as the competitions progressed, an increase of the Salivette's® contact time with the oral cavity along with the progression of a ride. How the level of salivary free cortisol is affected by reduced saliva warrants investigation [25]. High and low flow rates in normal adult humans did not show a difference in concentration in SC [51]. Even though the sample was smaller in this study, the highest increases of SC concentration still occurred in VG1, when horses were supposedly less dehydrated, and not in VG3.

It was also noticed that many saliva samples after extraction were contaminated with food. To investigate possible interference with the results a small trial was performed in five horses after a mouth wash to compare clean saliva and saliva posteriorly contaminated and incubated with different types of food (hay, granulated and grass). No significant differences ($p < 0.05$) were found between the different samples (MM et al. 2019, unpublished data). A recent study also showed that food contamination did not alter SC levels significantly [52].

Non-controlled interferences with IRT^{ET} The same operator recorded IRT^{ET} measurements during the research study and distance from the operator to the eye was measured at all times. However, we recognise that our values might have been affected by the environmental conditions' interference throughout the day. Ambient conditions, surface moisture, brightness, sun reflection and wind breeze are some of the variables that have been reported to interfere with IRT shooting [10]. A controlled environment as recommended [53] is challenging to achieve in endurance competitions, without interfering with the competition's pace and time management of the competitors.

Other parameters not quantified

Even if disrupted, how circadian rhythm could have influenced the variations from PI into VG1 was not taken into account. Also, the impact of different transportation times and characteristics, even over short-distances, could impact basal SC and IRT^{ET} was not quantified during travelling. Trainers were not questioned about the previous training of their horses. Prior competition history, including completions/eliminations rates and previous speed/recovery times/position records, were not analysed.

Measurement of body temperature was not included in this study due to its perceived invasiveness and practicality in young horses and less experienced horses participating in qualifier rides. Soroko et al. (2016) did not find a post-exercise correlation between rectal temperature and maximum eye temperature in 19 racehorses.

The rider's riding skills and weight might affect the emotional or physical response, e.g., the horse's capability to cope more or less efficiently with the demanded effort. In one study, SC was not affected by the rider's experience in showjumping [54], but the weight was found to have a detrimental effect on equine gait and behaviour. Furthermore, many times in endurance, the competitor might be just the pilot for that competition and not the rider training the horse, and whether this causes more stress to the horse was not studied before.

Future directions

It is still challenging to untangle emotional distress and experienced pain from the natural physiologic response to the effort. To exhaust the topic usefulness of biomarkers in identifying horses at risk during endurance competitions, more extensive studies are needed at high-level competitions, to collect statistically significant samples of horses that failed to qualify.

Conclusion

Pre-exercise baseline IRT^{ET} levels, but not SC, were higher in less experienced horses in the 40K compared to their counterparts in the 80K competitions. SC and IRT^{ET} showed different indications according to competition. In the 40 K competition, higher baseline pre-exercise

SC levels seemed to be linked to a better classification outcome. In contrast, in the 80 K horses, the higher IRTET variation from pre-exercise into final Vet Gate was the parameter associated with a better performance. A more controlled environment and a larger sample are needed to confirm these results and monitor horse welfare in competitions.

Methods

Horses

After competitors, owners and trainers were notified of the aims and methods of the research study to ensure informed consent, a convenience sample of 61 out of 110 horses participating in two endurance events in Portugal was obtained. Age was between 6 and 11 years and 24 were geldings, 29 were mares and eight were entire males. Breed varied; 30 horses were registered as Arabian, 27 as Anglo-Arabian or Part-Arabian and four as other breeds (one Lusitano and three from undetermined origin) in the Portuguese National Federation online database (www.fep.pt). All horses were transported the same day to the competition sites and travelling times were estimated according to their training stables' location.

Sampling moments and competition's features

Horses were sampled at home and at two competition sites at different times of the same year, e.g. June and November. At the Polo da Mitra of the University of Evora (MI) competition site, the saliva for cortisol determination of 23 horses was sampled. Of those, 14 horses were collected at their stables (Home), 22 to 24 h before the start of the event, depending on the owner or trainer's availability. At Torre de Palma Resort in Monforte (TP), 38 horses were sampled for cortisol and eye temperature was measured. Of the 61 horses that entered the study, 34 and 25 horses were participating in 40, and 80 km controlled speed (up to 16 km/h) qualifier rides, respectively, and only two horses were in an 80 Km free speed competition. For data processing competitors were grouped under 40 and 80K categories only.

The sampling took place following the veterinary inspections at the pre-inspection (PI) and upon completion of each phase immediately after the horses exited the Vet Gate area (VG), which was outdoor. If the horse failed to meet the heart rate criteria of 64 bpm, the collections were made after the heart rate reinspection. Requested or compulsory reinspections data were not used. The PI commenced at both sites at 7:00 AM and starts into the track took place in a staggered manner from 8:00 AM for the 80 km and 9:00 AM for the 40 km rides. The competitions finished around 3:00 PM.

Both 40 km qualifier rides were composed of two phases of 20 km, being the cumulative distance at VG1 20 km (VG1@20 km) and at VG2 40 km (VG2@40km). The 80 km rides had three phases, but with a different configuration between competition sites. At MI (80 K-A), the first phase had 40 km (VG1@40 km) and the remaining two had 20 km (VG2@60km and VG3@80km). At TP (80 K-B), there were two phases of 30 km (VG1@30km and VG2@60km) and only the last phase of 20 km (VG3@80km). Of the 61 horses that entered the study, 34 were competing in 40 km and 27 in 80 km rides.

Collection of saliva

A Salivette® (Starsted) synthetic swab was held on a metal clamp and maintained in every participating horse's mouth for 30–40 s, over and under the tongue, as described [28], and then placed into the Salivette® (Starsted) tube to be stored at 4 °C, at each collection moment. At the end of each day, the Salivettes were centrifuged for 10 min at 1500 *g* for saliva extraction and stored at − 28 °C until assayed. After thawing the samples, free cortisol was determined using a double-antibody immunoassay kit (Cortisol ELISA, IBL International GMBH, Germany).

Infrared thermographic eye temperature (IRT^{ET})

Eye temperature was measured using a portable infrared thermography camera (Thermal Imaging Camera, E60BX, FLIR Systems AB, Sweden) with 320 × 240 pixels set to emissivity 0.98. The sampling was performed outdoors in an open field after the horses exited the Vet Gate and before the saliva collection. To calibrate the camera results, environmental air temperature, and relative humidity were measured with a digital thermohygrometer (MR77, FLIR systems AB) at each collection. The left eye was scanned at a 90° angle at a distance of 1 m, as described previously [33], and several images were obtained. After electing the most adequate picture, an image analysis software (ThermaCam Researcher Pro, FLIR systems AB) was used to measure the maximal temperature within an oval area traced around the inner canthus of the eye, including the lacrimal caruncle at ~ 1 cm around the outside of the eyelids [55].

Performance and outcome data

Outcome and performance data (speed, recovery time and classification) were obtained from the veterinary cards and timing system. For analysis purposes, groups were created according to final position: Top5 (1 to 5th), G2 (from 6th to 10th) and G3 (from 11th). Those that failed to qualify were grouped under FTQ.

Data analysis

SPSS® version 22 software (Armonk, NY: IBM Corp.) was used for descriptive analysis and inferential statistics.

Variations in SC (ΔSC) were calculated as the percent of variation from one moment of collection to the following, according to the following formula:

$$\Delta SC = \frac{Mean\ SC\ (t+1) - Mean\ SC(t)}{Mean\ SC(t)}\ x\ 100$$

Where t is a determined moment of collection and t + 1 the following moment of collection. Variations in IRT^{ET} (ΔIRT^{ET}) were calculated in the same manner. Since data did not assume a normal distribution, a series of Kruskal Wallis analyses with post hoc Mann Whitney U tests identified significant differences between the variables recorded across ride categories, site, breed, and gender. A Wilcoxon signed-rank test assessed if cortisol and IRT^{ET} or variation of each measure were significantly different between collection moments. Where significance was found, post-hoc Bonferroni t-tests were used for multiple pairwise comparisons. A series of Spearman Rank Order correlations analysed if cortisol and IRT^{ET} were impacted by age and gender, speed and classification of the horses. Analysis significance was set at $P < 0.05$.

Abbreviations
40K: 40 km endurance ride; 80 K: 80 km endurance ride; **Δ**: Variation; FEI: Fédération Equestre Internationale; FTQ: Fail to Qualify; G2: Horses that were placed from 6th to 10th position; G3: Horses that were placed from 11th to 15th position; HPA: Hypothalamic-pituitary-adrenal; IRT^{ET}: Eye temperature measured by infrared thermography; MI: Polo da Mitra of the University of Evora competition site; PI: Pre-Inspection; ROC: receiver operating characteristic curve analysis; SC: Salivary cortisol; Top5: Horses that were placed in the first five positions; TP: Torre de Palma Resort competition site; VG: Vet Gate; VG1@20km: First vet gate after 20 km in competition; VG1@30km: First vet gate after 30 km in competition; VG1@40km: First vet gate after 40 km in competition; VG2@40km: Second vet gate after 40 km in competition; VG2@60km: Second vet gate after a 60 km in competition; VG3@80km: Third vet gate after a 80 km in competition

Acknowledgements
The authors gratefully acknowledge the Torre de Palma Resort organizing committee and the Portuguese Equestrian Federation officials working at the competitions for their support. The authors also acknowledge the FCT–Portuguese Science Foundation [grant UIDB/05064/2020] for author Rute Santos and for author Elsa Lamy through research contract CEECIND/04397/2017

Authors' contributions
MCM, RS and EL conceived and designed the study. MCM, MVP and PSM collected the data. EL AND PR were involved in the chemical analysis. MCM, RS, EL, JW and DM were involved in the data interpretation. MCM drafted the manuscript and RS, EL, JW and DM critically read and edited the manuscript. All authors read and approved the final manuscript.

Authors' information
This research is part of the PhD thesis under the title "Usefulness of non-invasive and objective methods in the assessment of the welfare of horses in endurance competitions" developed by the first author.

Funding
This work is funded by National Funds through FCT – Foundation for Science and Technology under the Project UIDB/05183/2020.

Availability of data and materials
The datasets are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
The department of animal welfare of the Portuguese Directorate-General for Food and Veterinary Affairs with the number 0421/000/000/2016 approved this project. Written owner's consent was obtained for all horses participating in this study.

Consent for publication
Not applicable.

Competing interests
Elsa Lamy is an editorial board member of BMC Veterinary Research, as associate editor. None of the other authors has any financial or personal relationships that could inappropriately influence or bias the content of the paper.

Author details
[1]MED – Mediterranean Institute for Agriculture, Environment and Development, Institute for Advanced Studies and Research, Universidade de Évora, Pólo da Mitra, Ap. 94, 7006-554 Évora, Portugal. [2]VALORIZA – Research Centre for Endogenous Resource Valorization, Edifício BioBIP, Campus Politécnico, 10, 7300-555 Portalegre, Portugal. [3]Hartpury University, Gloucester GL19 3BE, UK. [4]Departamento de Medicina Veterinária, Universidade de Évora, Núcleo da Mitra, Apartado, 94 7006-554 Évora, Portugal. [5]Departamento de Medicina Veterinária Portugal, Escola Universitária Vasco da Gama, Av. José R. Sousa Fernandes, Campus Universitário – Bloco B, Lordemão, 3020-210 Coimbra, Portugal. [6]David Marlin Consulting, AnimalWeb Ltd, Cambridge CB4 0WZ, UK.

Received: 16 November 2020 Accepted: 30 July 2021
Published online: 14 October 2021

References
1. Endurance Rules [https://inside.fei.org/sites/default/files/FEI%20Endurance%20Rules%20-%201%20July%202020%20-%2016.12.2019%20-%20Clean.pdf]. Accessed 15 Nov 2020.
2. Veterinary Regulations [https://inside.fei.org/fei/regulations/veterinary].
3. Nagy A, Murray JK, Dyson SJ. Descriptive epidemiology and risk factors for eliminations from federation Equestre Internationale endurance rides due to lameness and metabolic reasons (2008-2011). Equine Vet J. 2014;46(1):38–44. https://doi.org/10.1111/evj.12069.
4. Bennet ED, Parkin TDH. Fédération Equestre Internationale endurance events: risk factors for failure to qualify outcomes at the level of the horse, ride and rider (2010–2015). Vet J. 2018;236:44–8. https://doi.org/10.1016/j.tvjl.2018.04.011.
5. Marlin D, Williams J. Equine endurance race pacing strategy differs between finishers and non-finishers in 120 km single-day races. Comp Exercise Physiol. 2018;14(1):11–8. https://doi.org/10.3920/CEP170027.
6. Heleski C, Stowe C, Fiedler J, Peterson M, Brady C, Wickens C, et al. Thoroughbred Racehorse Welfare through the Lens of 'Social License to Operate—With an Emphasis on a U.S. Perspective. Sustainability. 2020;12(5):1706. https://doi.org/10.3390/su12051706.
7. Williams J, Marlin DJ. Foreword: emerging issues in equestrian practice. Comp Exercise Physiol. 2020;16(1):1–4. https://doi.org/10.3920/CEP20x001.
8. de Mira MC, Santos C, Lopes MA, Marlin DJ. Challenges encountered by federation Equestre Internationale (FEI) veterinarians in gait evaluation during FEI endurance competitions: an international survey. Comp Exercise Physiol. 2019;15(5):371–8. https://doi.org/10.3920/CEP180058.
9. van Loon JPAM, Van Dierendonck MC. Objective pain assessment in horses (2014–2018). Vet J. 2018;242:1–7. https://doi.org/10.1016/j.tvjl.2018.10.001.
10. König V, Borstel U, Visser EK, Hall C. Indicators of stress in equitation. Appl Anim Behav Sci. 2017;190:43–56. https://doi.org/10.1016/j.applanim.2017.02.018.
11. Bartolomé E, Cockram MS. Potential effects of stress on the performance of sport horses. J Equine Vet Sci. 2016;40:84–93. https://doi.org/10.1016/j.jevs.2016.01.016.

12. Schmidt A, Möstl E, Wehnert C, Aurich J, Müller J, Aurich C. Cortisol release and heart rate variability in horses during road transport. Horm Behav. 2010; 57(2):209–15. https://doi.org/10.1016/j.yhbeh.2009.11.003.
13. Becker-Birck M, Schmidt A, Lasarzik J, Aurich J, Möstl E, Aurich C. Cortisol release and heart rate variability in sport horses participating in equestrian competitions. J Vet Behav. 2013;8(2):87–94. https://doi.org/10.1016/j.jveb.2012.05.002.
14. Hall C, Kay R, Yarnell K. Assessing ridden horse behavior: professional judgment and physiological measures. J Vet Behav. 2014;9(1):22–9. https://doi.org/10.1016/j.jveb.2013.09.005.
15. Peeters M, Closson C, Beckers J-F, Vandenheede M. Rider and horse salivary cortisol levels during competition and impact on performance. J Equine Vet Sci. 2013;33(3):155–60. https://doi.org/10.1016/j.jevs.2012.05.073.
16. Hartmann E, Christensen JW, Keeling LJ. Training young horses to social separation: effect of a companion horse on training efficiency. Equine Vet J. 2011;43(5):580–4. https://doi.org/10.1111/j.2042-3306.2010.00326.x.
17. Dyson S, Berger JM, Ellis AD, Mullard J. Development of an ethogram for a pain scoring system in ridden horses and its application to determine the presence of musculoskeletal pain. J Vet Behav. 2018;23:4W 57.
18. Hyyppä S. Endocrinal responses in exercising horses. Livest Prod Sci. 2005; 92(2):113–21. https://doi.org/10.1016/j.livprodsci.2004.11.014.
19. Peeters M, Sulon J, Beckers JF, Ledoux D, Vandenheede M. Comparison between blood serum and salivary cortisol concentrations in horses using an adrenocorticotropic hormone challenge. Equine Vet J. 2011;43(4):487–93. https://doi.org/10.1111/j.2042-3306.2010.00294.x.
20. Janczarek I, Bereznowski A, Strzelec K. The influence of selected factors and sport results of endurance horses on their saliva cortisol concentration. Pol J Vet Sci. 2013;16(3):533–41. https://doi.org/10.2478/pjvs-2013-0074.
21. Rose RJ, Hodgson DR, Sampson D, Chan W. Changes in plasma biochemistry in horses competing in a 160 km endurance ride. Aust Vet J. 1983;60(4):101–5. https://doi.org/10.1111/j.1751-0813.1983.tb05905.x.
22. Kędzierski W, Cywińska A. The effect of different physical exercise on plasma leptin, cortisol, and some energetic parameters concentrations in purebred Arabian horses. J Equine Vet Sci. 2014;34(9):1059–63. https://doi.org/10.1016/j.jevs.2014.06.005.
23. Jastrzębska E, Wolska A, Minero M, Ogłuszka M, Earley B, Wejer J, et al. Conflict behavior in show jumping horses: a field study. J Equine Vet Sci. 2017;57:116–21. https://doi.org/10.1016/j.jevs.2017.07.009.
24. Cayado P, Muñoz-Escassi B, Domínguez C, Manley W, Olabarri B, Sánchez de la Muela M, et al. Hormone response to training and competition in athletic horses. Equine Vet J Suppl. 2006;36:274–8.
25. Bohak Z, Szabo F, Beckers JF, Melo de Sousa N, Kutasi O, Nagy K, et al. Monitoring the circadian rhythm of serum and salivary cortisol concentrations in the horse. Domest Anim Endocrinol. 2013;45(1):38–42. https://doi.org/10.1016/j.domaniend.2013.04.001.
26. Hellhammer DH, Wust S, Kudielka BM. Salivary cortisol as a biomarker in stress research. Psychoneuroendocrinology. 2009;34(2):163–71. https://doi.org/10.1016/j.psyneuen.2008.10.026.
27. Duclos M, Corcuff JB, Arsac L, Moreau-Gaudry F, Rashedi M, Roger P, et al. Corticotroph axis sensitivity after exercise in endurance-trained athletes. Clin Endocrinol. 1998;48(4):493–501. https://doi.org/10.1046/j.1365-2265.1998.00334.x.
28. Peeters M, Sulon J, Serteyn DA, Vandenheede M. Assessment of stress level in horses during competition using salivary cortisol: preliminary studies. J Vet Behav. 2010;5(4):216. https://doi.org/10.1016/j.jveb.2009.10.043.
29. Munk R, Jensen RB, Palme R, Munksgaard L, Christensen JW. An exploratory study of competition scores and salivary cortisol concentrations in warmblood horses. Domest Anim Endocrinol. 2017;61:108–16. https://doi.org/10.1016/j.domaniend.2017.06.007.
30. Okada K, Takemura K, Sato S. Investigation of various essential factors for optimum infrared thermography. J Vet Med Sci. 2013;75(10):1349–53. https://doi.org/10.1292/jvms.13-0133.
31. Esteves Trindade PH, de Camargo FG, Pereira Lima ML, Negrão JA, Paranhos da Costa MJR. Eye surface temperature as a potential Indicator of physical fitness in ranch horses. J Equine Vet Sci. 2019;75:1–8. https://doi.org/10.1016/j.jevs.2018.11.015.
32. Soroko M, Howell K, Zwyrzykowska A, Dudek K, Zielińska P, Kupczyński R. Maximum eye temperature in the assessment of training in racehorses: correlations with salivary cortisol concentration, rectal temperature, and heart rate. J Equine Vet Sci. 2016;45:39–45. https://doi.org/10.1016/j.jevs.2016.06.005.
33. Valera M, Bartolomé E, Sánchez MJ, Molina A, Cook N, Schaefer A. Changes in eye temperature and stress assessment in horses during show jumping competitions. J Equine Vet Sci. 2012;32(12):827–30. https://doi.org/10.1016/j.jevs.2012.03.005.
34. Yarnell K, Hall C, Billett E. An assessment of the aversive nature of an animal management procedure (clipping) using behavioral and physiological measures. Physiol Behav. 2013;118:32–9. https://doi.org/10.1016/j.physbeh.2013.05.013.
35. Negro S, Bartolomé E, Molina A, Solé M, Gómez MD, Valera M. Stress level effects on sport performance during trotting races in Spanish trotter horses. Res Vet Sci. 2018;118:86–90. https://doi.org/10.1016/j.rvsc.2018.01.017.
36. Sánchez MJ, Bartolomé E, Valera M. Genetic study of stress assessed with infrared thermography during dressage competitions in the Pura Raza Español horse. Appl Anim Behav Sci. 2016;174:58–65. https://doi.org/10.1016/j.applanim.2015.11.006.
37. Irvine CHG, Alexander SL. Factors affecting the circadian rhythm in plasma cortisol concentrations in the horse. Domest Anim Endocrinol. 1994;11(2): 227–38. https://doi.org/10.1016/0739-7240(94)90030-2.
38. Alexander S, Irvine C. The effect of social stress on adrenal axis activity in horses: the importance of monitoring corticosteroid-binding globulin capacity. J Endocrinol. 1998;157(3):425–32. https://doi.org/10.1677/joe.0.1570425.
39. Berghold P, Möstl E, Aurich C. Effects of reproductive status and management on cortisol secretion and fertility of oestrous horse mares. Anim Reprod Sci. 2007;102(3):276–85. https://doi.org/10.1016/j.anireprosci.2006.11.009.
40. Fazio E, Medica P, Cravana C, Ferlazzo A. Effects of competition experience and transportation on the adrenocortical and thyroid responses of horses. Vet Rec. 2008;163(24):713–6. https://doi.org/10.1136/vr.163.24.713.
41. de Graaf-Roelfsema E, Keizer HA, van Breda E, Wijnberg ID, van der Kolk JH. Hormonal responses to acute exercise, training and overtraining. A review with emphasis on the horse. Vet Q. 2007;29(3):82–101. https://doi.org/10.1080/01652176.2007.9695232.
42. Viru A, Viru M. Cortisol--essential adaptation hormone in exercise. Int J Sports Med. 2004;25(6):461–4. https://doi.org/10.1055/s-2004-821068.
43. Pedersen BK, Steensberg A, Schjerling P. Muscle-derived interleukin-6: possible biological effects. J Physiol. 2001;536(Pt 2):329–37. https://doi.org/10.1111/j.1469-7793.2001.0329c.xd.
44. Viru AM, Hackney AC, Välja E, Karelson K, Janson T, Viru M. Influence of prolonged continuous exercise on hormone responses to subsequent exercise in humans. Eur J Appl Physiol. 2001;85(6):578–85. https://doi.org/10.1007/s004210100498.
45. Fenner K, Yoon S, White P, Starling M, McGreevy P. The effect of noseband tightening on Horses' behavior, eye temperature, and cardiac responses. PLoS One. 2016;11(5):e0154179. https://doi.org/10.1371/journal.pone.0154179.
46. Bartolome E, Sanchez MJ, Molina A, Schaefer AL, Cervantes I, Valera M. Using eye temperature and heart rate for stress assessment in young horses competing in jumping competitions and its possible influence on sport performance. Animal. 2013;7(12):2044–53. https://doi.org/10.1017/S1751731113001626.
47. Redaelli V, Luzi F, Mazzola S, Bariffi GD, Zappaterra M, Nanni Costa L, et al. The Use of Infrared Thermography (IRT) as Stress Indicator in Horses Trained for Endurance: A Pilot Study. Animals (Basel). 2019;9(3):84.
48. Cook N, Schaefer A, Warren L, Burwash L, Anderson M, Baron V. Adrenocortical and metabolic responses to ACTH injection in horses: an assessment by salivary cortisol and infrared thermography of the eye. Can J Anim Sci. 2001;81:621.
49. Anderson T, Wideman L. Exercise and the cortisol awakening response: a systematic review. Sports Med Open. 2017;3(1):37. https://doi.org/10.1186/s40798-017-0102-3.
50. Mircean M, Giurgiu G, Mircean V, Zinveliu E: Serum cortisol variation of sport horses in relation with the level of training and effort intensity. Bull Univ Agric Sci. 2008:64(1–2). http://journals.usamvcluj.ro/index.php/veterinary/article/view/2481. Accessed Apr 2020.
51. Vining RF, McGinley RA. Hormones in saliva. Crit Rev Clin Lab Sci. 1986;23(2): 95–146. https://doi.org/10.3109/10408368609165797.
52. Contreras-Aguilar MD, Hevia ML, Escribano D, Lamy E, Tecles F, Cerón JJ. Effect of food contamination and collection material in the measurement of biomarkers in saliva of horses. Res Vet Sci. 2020;129:90–5. https://doi.org/10.1016/j.rvsc.2020.01.006.

53. Purohit RC: Standards for thermal imaging in veterinary medicine. In: Proceedings of the 11th European Congress of Medical Thermology. 2009; 99. https://doi.org/10.13140/RG.2.1.2015.1206.
54. Ille N, Lewinski M, Erber R, Wulf M, Aurich J, Möstl E, et al. Effects of the level of experience of horses and their riders on cortisol release, heart rate and heart-rate variability during a jumping course. Anim Welf. 2013;22(4): 457–65. https://doi.org/10.7120/09627286.22.4.457.
55. Dai F, Cogi NH, Heinzl EUL, Dalla Costa E, Canali E, Minero M. Validation of a fear test in sport horses using infrared thermography. J Vet Behav. 2015; 10(2):128–36. https://doi.org/10.1016/j.jveb.2014.12.001.

Publisher's Note

CHAPTER VII: General Discussion and Conclusions

The welfare of horses and objectiveness of human decisions towards animals is increasingly a major concern as we progress into the XXI century. This work was devoted to exploring methods that could be potentially used during equestrian competitions for their non-invasiveness to improve equine athletes' welfare monitoring and competitor's equity. Specifically in endurance, recognized as the most physiologically challenging equestrian discipline for the horse and therefore intensively monitored, horses are often failed to qualify by veterinarians, whose judgment relies mainly on subjective parameters. Especially gait assessment that accounts for two-thirds of the horses being failed to continue in competition or qualify is increasingly questioned by competitors, for all in a high technology era, where trot-ups are easily registered anytime by cameras. Subjective assessment can generate type I and II errors, e.g. failing horses that could complete the competition and not failing horses that should have been failed. False negatives will impact directly on horses' welfare and health, while false positives impact mainly on competitors, trainers, and owners. Therefore, in the absence of objective measures, judgments will mostly err on errors type I, since horse morbidity, or even mortality, self-justifies the unbalance since welfare must be protected at all costs. However, there is a cost also for horses. First, premature retirement and wastage of horses that could still compete and make a career in endurance, as their counterparts do in other disciplines, much more tolerant towards gait irregularities. Secondly, a quest for illicit measures to prevent by any means that a horse goes lame during competition, as most competitors see gait assessment results as unpredictable, not only from the horse's point of view that can always suffer an injury but also regarding veterinarians' judgment.

In contrast to gait evaluation, metabolic injuries risk assessment relies on more objective parameters, such as heart rate and recovery times, much more intuitive for competitors, generating therefore much less opposition when resulting in failure to qualify.

This work, a preliminary study on non-invasive objective methods that could be used in endurance competitions, explored necessity, implementation feasibility, and preliminary results regarding the use of a sensor-based method to quantify gait, salivary cortisol, and infra-red thermography. It also provided useful evidence-based information that can improve guidelines in rules of ruling organisms such as the FEI, including better training of officials and competitors.

1. Necessity of objective measurements in endurance competitions

This work identified significant limitations, previously reported only anecdotally, of subjective gait assessment during competitions. This evidence-based information is an essential reference for further research in gait analysis **and** riding endurance sport.

First of all, it demonstrated that identification of lameness during competitions is not always straightforward for most veterinarians (de Mira M.C. *et al.*, 2019). Furthermore, it identified clearly that handling of the horse is the main factor interfering with a clear judgement, already limited by the logistics' conditioning factors inherent to a competition. Therefore, this topic was extensively analysed in the experimental study "Vet gate trotting style improves subjective gait grading in endurance horses when compared to a lameness presentation style (chapter V)", showing that indeed presentation trotting styles coupled with speed can interfere with gait assessment inducing a lower scoring in lameness classification.

Secondly, the interest of veterinarians towards the use of technologies to quantify gait in endurance competitions was demonstrated, with less than one-third of the 157 surveyed FEI officials refuting its use (de Mira M.C. *et al.*, 2019). Interestingly, a recent study performed with a more generalist population of veterinarians produced similar results (Hardeman *et al.*, 2021). Moreover, competitors showed receptivity to a sensor-based quantifying system even in high-level competitions after a trial in qualifying endurance rides (Lopes, 2018).

2. Implementation feasibility

The trial using a sensor-based gait quantifying system performed in the vet gates of qualifying endurance rides competitions showed its feasibility in endurance competitions (Lopes, 2018). Furthermore, the study showed that only one back-and-forth trot-up is necessary for the objective readings, even if the 25 recommended strides are not attained.

Moreover, the potential interference with measuring the two heart rates needed to calculate the Cardiac Recovery Index (CRI) was demonstrated to be innocuous (de Mira *et al.*, 2020). Still, even if validated to be used in a medium term in competition settings, the cost and personnel needed to instrument the horses make the use of existent gait quantifying systems prohibitive.

3. Preliminary results

Significant preliminary results of using non-invasive objective methods to quantify gait (Lopes, 2018) and stress (de Mira *et al.*, 2021) were presented in this work. In a first moment, it was shown that thresholds in use to detect clinical asymmetry with sensor-based systems are too strict to be used in competitions. Only when thresholds were doubled, an agreement could be found between the veterinarian's subjective and the objective assessment. However, this could also be explained by the human eye's lower sensitivity, especially to horses trotted at competitions, since handlers will make every effort to hide any potential lameness at a vet gate. Actually, and probably the most crucial result presented by this work is the demonstration that a vet gate trot-up presentation style coupled with increased speed can lower the subjective lameness scores. However, preliminary results (unpublished) have shown that this is not the case with objective gait analysis.

Salivary cortisol and infrared thermography eye temperature are known to stress biomarkers that will naturally increase with effort and are difficult to distinguish from competition-induced morbidities. Unfortunately, we did not have a large enough sample to compare finishers with horses failing to qualify to assess a significant difference. Nevertheless, we could show that starting a competition with high cortisol levels in the most inexperienced group might enhance performance. Furthermore, a higher eye temperature variation might be indicative of endurance ability in horses.

4. Guidelines' improvement

This work might improve the governing bodies' guidelines, namely the FEI, regarding the sport's ruling. One example was the results of our study (de Mira *et al.*, 2020) that brought to the FEI attention that the guidelines given to veterinarians of how to perform the procedure (FEI, 2019) did not match the original explanation (Ridgeway, 1991). However, we showed evidence that the original form elicited less variability in the results, and the rules were changed accordingly (FEI, 2020).

Other important contributes:

- Caution not to over-interpret the CRI when the first heart rate is lower than 60 bpm or closer to baseline as in re-inspections, since a rise more than 4bpm might be physiological should be transmitted to endurance veterinarians (de Mira *et al.*, 2020)

- The impact of trot-ups presentation to the already challenging task of gait assessment in competitions should be addressed in the rules with clear guidelines to the handlers and videoed based training material for all involved in the competition

- Pointing to the future of the mid-term role of objective gait analysis in competitions and the necessary FEI support for studies in real setting competitions.

References:

de Mira M.C., Santos, C., Lopes, M.A. and Marlin, D.J., 2019. Challenges encountered by Federation Equestre Internationale (FEI) veterinarians in gait evaluation during FEI endurance competitions: an international survey. Comparative Exercise Physiology 15: 371-378. https://doi.org/10.3920/cep180058

de Mira, M.C., Lamy, E., Santos, R.., Williams, J., Vaz Pinto, M.., Martins, P., Rodrigues, P., and Marlin, D., Pre-print. Salivary Cortisol and Infrared Thermographic Ocular Temperature Use as Biomarkers during Endurance Competitions. BMC Veterinary Research. https//:10.21203/rs.3.rs-111435/v1

de Mira, M.C., Williams, J., Santos, C., Rodrigues, P., Arroja, B. and Marlin, D., 2020. Do waiting times in endurance vet gates affect the cardiac recovery index? Comparative Exercise Physiology: 1-8. https://doi.org/10.3920/CEP190081

FEI, 2019. Veterinary Regulations. Available at: https://inside.fei.org/sites/default/files/FEI%20Rules/FEI%20Veterinary%20Regulations%202020%20-%20Clean%20Version.pdf.

FEI, 2020. Veterinary Regulations. Fédération Equestre Internationale, Lausanne, Switzerland.

Hardeman, A.M., Van Weeren, P.R., Serra Bragança, F.M., Warmerdam, H. and Bok, H.G.J., 2021. A first exploration of perceived pros and cons of quantitative gait analysis in equine clinical practice. Equine Veterinary Education n/a. https://doi.org/10.1111/eve.13505

Lopes, M., de Mira, MC, Santos, C, 2018. Feasibility of Objective Detection and Quantification of Gait Abnormalities with a Portable Inertial Sensor-based System in Horses During an Endurance Race.

Ridgeway, K.J., 1991. Inride veterinary examination, postride examination and judging of best condition., American Association of Equine Practioners Annual Convention (AAEP). pp. 815-826.

ANNEXES

Annex I – Ethics approval and consent to participate. The department of animal welfare of the Portuguese Directorate-General for Food and Veterinary Affairs with the number 0421/000/000/2016 approved this project. Written owner's consent was obtained for all horses participating in this study.

REPÚBLICA PORTUGUESA | AGRICULTURA, FLORESTAS E DESENVOLVIMENTO RURAL

dgav Direção Geral de Alimentação e Veterinária

Ex^ma^ Senhora
Drª Mónica Mira
Universidade de Évora
Pólo da Mitra, Apartado 94
7002 – 554 ÉVORA

2016-11-15 022082

Nossa referência	Vossa referência	Vossa data
0421/000/000 /2016		

Assunto: **PROTEÇÃO DOS ANIMAIS UTILIZADOS PARA FINS EXPERIMENTAIS E/OU OUTROS FINS CIENTÍFICOS – PEDIDO DE AUTORIZAÇÃO PARA REALIZAÇÃO DE PROJECTO DE EXPERIMENTAÇÃO ANIMAL**

Na sequência do pedido efetuado por V. Exª no sentido de poder ser autorizada a realização do projeto experimental designado **"Application of objective and non-invasive methodologies for assessment of welfare in endurance horses"**, de que é a investigadora responsável, cabe-me informar que o mesmo foi avaliado de acordo com o Artigo 44º do Decreto-Lei nº 113/2013, de 7 de Agosto, relativo à "proteção dos animais utilizados para fins científicos".

Mais se informa V. Exª que, depois de esclarecidas as dúvidas que a sua análise nos levantou, o projeto em apreço recebeu uma **avaliação favorável** e foi **autorizado** de acordo com o nº 1, do Artigo 42º do mesmo diploma legislativo.

No entanto, e no que diz respeito à severidade dos procedimentos deverá, fazer-se para cada um dos procedimentos a realizar, uma adequada monitorização dos sinais de dor, sofrimento ou angústia dos animais neles envolvidos, por forma a poder fazer-se uma atualização sobre o nível de dor efetiva a que os mesmos possam vir a ficar sujeitos independentemente de os mesmos serem classificados com a categoria **Ligeiro**.

Finalmente, resta-me especificar, de acordo com o discriminado no nº 2, do Artigo 46º, do atrás referido Decreto-Lei, o seguinte:

- **A pessoa responsável pela execução global do projeto e pela sua conformidade com a autorização do mesmo:** Drª Mónica Mira;

Com os melhores cumprimentos,

O Diretor Geral

Fernando Bernardo

Graça Mariano
Subdiretora Geral

DBEA/APM

SEDE : CAMPO GRANDE, 50 – 1700-093 LISBOA - TELEF. 21 323 95 00 FAX. 21 346 35 18

Annex II- FEI guidelines for endurance veterinarians as published in Bennet ED, Hayes ME, Friend L, Parkin TDH. The association between clinical parameters recorded at vet gates during Fédération Equestre Internationale endurance rides and the imminent risk of elimination. Equine Vet J. 2020 Nov;52(6):832-840. doi: 10.1111/evj.13264. Epub 2020 Apr 17. PMID: 32219883.

Equine Veterinary Journal BEVA

Text S1: Measurement of Vet Gate Parameters

Guidance and definitions from a course provided for FEI1* endurance vets (Drs Martha Misheff and Sarah Coombs personal communication). Note that these definitions come from experienced endurance vets and are not currently included in the FEI endurance rules.

Capillary Refill Time (CRT) 1-4
The number of seconds for colour to return to the mucous membranes of the gum after applying digital pressure to cause blanching. Counted as: one thousand one, one thousand two, etc.
1- One second- Normal
2- Two seconds- Normal capillary refill time should usually be less than two seconds.
3- A CRT of 3 seconds, with one or more additional parameters of concern, should be the impetus to call a Metabolic vote.
4- A CRT of 4 seconds should always stimulate a call for a vote, but this will rarely occur without other signs of distress.

Mucous Membranes (MM) A-D
A- Pink, moist
B- Pale pink/ less moist
C- Darkened/Discoloured /Tacky- A MM score of C, with one or more additional parameters of concern, should be the impetus to call a Metabolic vote.
D- Purple/Blue (cyanotic)- A MM score of D should always stimulate a call for a vote, but this will rarely occur without other signs of distress.

Gastrointestinal Motility (GI) A-D
A- Normal auscultable motility
B- Decreased auscultable motility
C- Poorly discernable auscultable motility
D- Absent auscultable motility

This parameter alone is not significant if the horse has presented in a short time. Blood is preferentially shunted to muscles during exercise, so many horses may have little or no ascultable motility when coming into the Horse Inspection Area directly from the course. On re-examination, after a rest period, or if the horse has taken 15-20 minutes to present initially, lack of ascultable GI motility is potentially more significant.

Hydration: 1-4
Counted in seconds: one thousand one, one thousand two, etc.

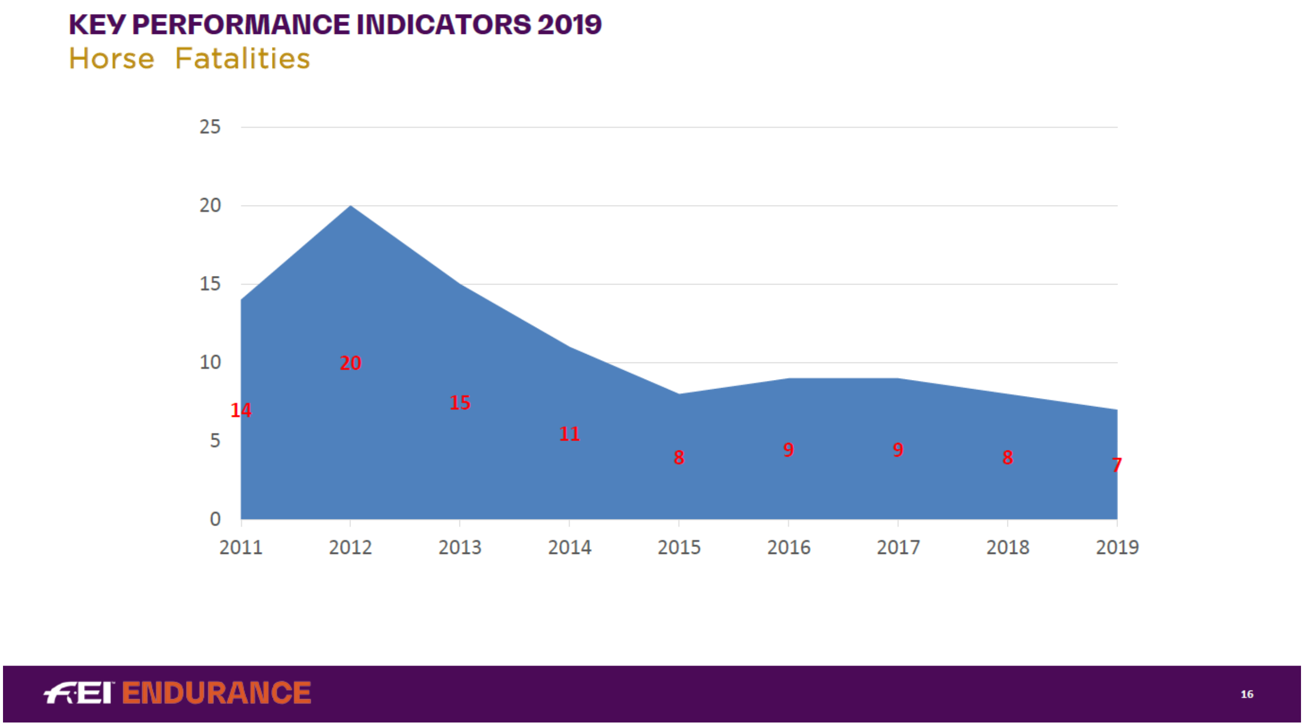
This parameter is the least reliable of all recorded metabolic indices, as skin turgour varies considerably amongst horses due to age and skin elasticity, as well as hydration status. It should be measured at the point of the shoulder, but variability persists in the measured location amongst veterinarians.

Girth, back and withers soreness: A-D
A- no sensitivity when palpated with the flat of the hand, and no sign of any lesions
B- some sensitivity or a lesion that needs closer inspection and flagging up to keep an eye on.
C/D- soreness or recent open lesion that needs voting and fails.

Gait: A-D
A- is ok. We don't say 'sound' because that suggests perfect, but on the basis of 'rhythm, elasticity and vigour' - and I would add symmetry- they don't give cause for concern and are fit to continue.
B- is passable but not as good, maybe the odd inconsistent stride or needs to be trotted for a second opinion but passed. Often can be indicative of a tired/slightly dehydrated or electro. depleted horse not showing itself off very well.
C- is clearly lame.
D- is severely lame.

Annex III – Fatalities in endurance competitions a.) Occurrence in general and b-) broken down by cause in FEI endurance competitions published at https://inside.fei.org/sites/default/files/FEI_Endurance_Report_2019.final.pdf

a.)



b.)