PhD thesis: 'Using data to track racehorse physiology during training and racing'

Charlotte Schrurs 14334875

Examiners' comments – April 2024

Notes: Any changes are conducted in blue throughout the thesis.

Page 16:
Please pay attention to British vs USA English particularly in words containing "z".
All verbs and words cross-checked through the advanced search function in MS Word and verified against the Cambridge Dictionary to ensure consistent British spelling.

Please avoid the use of 'trend' in the text.
All words identified through the advanced search function in MS Word and changed to the below:

- i: phrase revised to: 'There was clearly a zone of maximal cardiac flexibility, at hard canter-to-slow gallop, where responses to training were greatest, perhaps offering the opportunity to use as a basis for setting training zones, as opposed to maximal heart rate, which offered little in terms of between or within horse performance difference.'
- Page 35; replaced by 'function'. Treadmill research has expanded the understanding of physiological responses to exercise, encompassing cardiovascular, respiratory, metabolic, thermoregulatory, hematologic, and hormonal functions at various distances and intensities (Hodgson and Rose, 1994).
- Page 69: replaced by 'relationship': Previous studies have demonstrated that horses exhibit a linear relationship between heart rate and submaximal exercise intensity (Hinchcliff et al., 2013).
- Page 88: sentence rephrased: In contrast, the opposite was observed on sand; racehorses recovered more quickly with male riders, but only after medium-hard gallops (n=68 gallops on sand; Figure 9b,d).
- Page 108: replaced by 'small effect'. There was a small effect for racehorses with longer peak stride length (i.e., stayers) to have an increased chance of winning and finishing in the top 3, regardless of race class and racehorse type (**Error! Reference source not found.**b).
- Page 111: replaced by 'findings'. These findings were consistent with the ages observed across the three racehorse profiles (sprinter, miler, or stayer) in this study; younger horses tend to be sprinters, whilst milers and stayers were significantly older.
- Page 131: replaced by 'consistent but small effect'. Overall, a consistent but small effect was observed over the 5 first minutes following exercise: horses that had previously participated in a Group race recovered faster than those that hadn't.
- Page 131: replaced by 'showed similar measures of heart rate'. Both groups showed similar measures of heart rate up to the 15-minute recovery mark.
- Page 133: sentence rephrased. Stayers presented an increased chance of winning as opposed to the other horse profiles (i.e milers), with sprinters as the referent category.
- Page 140: replaced by 'small effect'. In addition, horses recovered better in relatively hot, but not relatively dry, weather conditions.

All figures and tables legends need careful revision to make them more reader friendly.
The legends describe the data recorded accurately and the description that is written has been accepted by peer-reviewers and considered as acceptable for publication. Unsure as to why we should make it more reader-friendly for a thesis.

Most descriptions are confusing and lack precision statements.

The legends just describe the data as presented in the figures so that the figures can be read on their own. As previously stated, since they were all accepted via peer-review, we believe that for a thesis they are satisfactory.

Make sure you add error bars to your figures when you say that data is displayed with it.

For many, the error bars are too small to be seen on the graph axis that shows the effect. That is because most of the error in the data has been accounted for and thus predicted means are accurate and sensitive.

Page 17:

How can your work help to minimise racehorse wastage?

Page 16 sentence added: To date, limited information is available to quantify this wastage effect. However, suitable training programs coupled to objective performance monitoring (i.e by means of validated tracking technologies) could lead to reduced wastage and overall improved health and welfare.

Confusing "thirds" how many TBs go finally on racing?

Page 16 phrasing altered: Within the horseracing discipline, and across other equine sports, a 'one-third' theory continues to be observed: within any given year cohort, approximately one-third of the horses born will not enter the racing sport. Within the remaining cohort, horses will either be intentionally withdrawn from training (e.g. unsuitability for the sport), retire unintentionally (e.g. poor performance-related issues such as musculoskeletal injuries) or will be sent to breeding farms.

Page 24: Clarify differences between sand and dirt.

I have clarified the difference on Page 24: Various surfaces compose racetracks, each with their own characteristics that can significantly affect the performance of horses. The main types of racetrack surfaces include dirt, turf, synthetic and sand. The choice of surface depends on factors such as climate, location, and the preferences of racing authorities.

Dirt tracks are commonly used in the United States and consist of a mixture of sand, clay, and soil. They are often treated with additives to improve both stability and drainage and provide a loose surface that becomes compact over time. Their offer a cushioning effect reducing impact on racehorses' limbs while remaining weather dependent (i.e., muddy with rain, dusty when dry).

Turf tracks are made of grass and are commonly used in Europe, Asia, and some parts of the United States. Turf racing is often associated with long distance races and is favoured by horses with a preference for grass surfaces. Turf tracks are generally firmer than dirt tracks, providing a consistent and stable surface. The disadvantage of turf tracks lies in the slow draining of the ground. Grass provides a softer surface, reducing the impact of racehorse running forces, which can be beneficial for horses' joints.

Also known as all-weather tracks, synthetic tracks are artificial surfaces designed to provide a consistent and stable surface, good drainage properties and resist to extreme weather conditions. Made from artificial materials such as polytrack, or other proprietary blends (e.g., wax or polymeric binder on a hard porous base), they replicate the properties of natural surfaces (e.g dirt and turf), providing a reliable racing surface that can withstand heavy rain or dry spells.

Some racetracks use sand tracks, consisting of a mixture of sand and other materials. Man-made sand tracks tend to be composed of pure sand over a non-porous base, to offer drainage and cushioning. When used in training, they provide a soft and 'forgiving' surface for conditioning work. In Australia for example, location dependent, it is not uncommon for trainers to exercise their horses on the beach on firm or soft sand (Morrice-West et al., 2018). Sand tracks are less common in major racing circuits but may be found in certain regions where other surfaces are not practical due to climate or cost considerations.

Page 33: Catastrophic injury: is age a risk factor or not? Conflicting information?
This is already explained on page 17: 'Contrasted opinions fuel the debate amongst equine professionals, as this practice engages relatively immature horses to high work demands early on in their development phase and may therefore increase the risk of injury. Evidence in the literature on both the training and racing of two-year-olds is conflicting.'
And on page 19: The following factors are believed to mitigate CMI occurring on the racetrack: horse age.

Are fatty acids the main fuel in racing? Please explain that section better.
Page 29: paragraph rephrased: The aerobic pathway relies on oxygen to break down nutrients, primarily fatty acids and carbohydrates, to produce energy in the form of adenosine triphosphate (ATP). In racehorses, the aerobic pathway is primarily responsible for providing sustained energy during moderate-intensity exercise, such as long-distance gallops and endurance rides. During aerobic metabolism, fatty acids are broken down in the mitochondria of muscle cells through a process called beta-oxidation, while carbohydrates are metabolized through glycolysis (anaerobic) and the citric acid cycle (aerobic). Full oxidation of hexose sugars like glucose to Co2 and H2O yields a large amount of ATP per molecule of nutrient, making it an efficient energy source for prolonged, low-moderate intensity activities.

While the anaerobic pathway operates in the absence of sufficient oxygen and is utilized during high-intensity, short-duration efforts, such as sprints or intense bursts of speed. In racehorses, the anaerobic pathway primarily relies on the rapid breakdown of stored carbohydrates (glycogen) in muscle cells to produce energy. Anaerobic metabolism involves glycolysis, where glucose is converted into pyruvate in the cytoplasm of cells, followed by the conversion of pyruvate into lactate or lactic acid. While the anaerobic pathway produces ATP quickly, it is less efficient than aerobic metabolism and generates lactate as a byproduct. Accumulation of lactic acid in muscles leads to difficulties in sustaining physical exertion, commonly referred to muscular fatigue or 'heaviness' in human physiology (Sahlin, 1986). The ability to maintain high-intensity exercise is thus associated with the capacity to either prevent the production of lactic acid or enhance its metabolism to minimise accumulation.

Page 36: Expand this section as account for your method of data collection, explain and show through pictures Equimetre, where sensors and electrode are.

For the Chapters 2-5, data collection occurred retrospectively for all the available parameters, as part of the company's database. The trackers were fitted to the horses by trained stable personnel and thus, data was directly collected by users of the Equimetre (i.e riders or racehorse trainers). As the individual studies developed over time, the parameters and algorithms of the company's database were constantly upgraded and refined. Therefore, separate databases were provided for each study.

The Equimetre comprises of two fixing systems, the first is an all-in-one system. One electrode is placed under the girth, close to the heart, and the other electrode is positioned under the saddle pad. The second system consists of a strap with two electrodes, one close to the heart and the other to the right of the withers. The girth is connected to the device unit secured at the right rear of the saddle pad by a wired cable. The electrodes are positioned as previously illustrated below, by ter Woort et al. (2021) (**Error! Reference source not found.**a).

How does this new technology compare to others?
Page 36 – Equine tracking tools subsection: I have included a comparative table, with permission from the corresponding author, on the existing commercially available equine

trackers excerpt from the following article: A synopsis of wearable commercially available biometric monitoring devices and their potential applications during gallop racing by Kee et al. (2023). Link to article: https://beva.onlinelibrary.wiley.com/doi/pdf/10.1111/eve.13800

Appendix on the device validation?
Page 42, paragraph added on the validation aspects of the Equimetre heart rate, locomotion and speed measurement.
The validity of the Equimetre for heart rate monitoring, thus ECG, was evaluated against the Televet system in 49 exercising horses and was concluded reliable including for arrythmia detection (ter Woort et al., 2021).

As far as locomotion is concerned, the company has integrated algorithms developed in the 90s by Barrey et al. (1995). There are also internal validations which the company could not disclose to include in this thesis, for obvious intellectual property reasons. Stride information is captured by gyroscope (i.e a device that measures orientation and angular velocity) and accelerometer sensors. The stride length calculation combines data from both sensors. The tracker can provide a stride length estimate by analysing the time taken for a full stride cycle and the associated acceleration/deceleration.

Speed measurements rely on a commercially available integrated system, combining GPS (USA), GLONASS (RUS) and GALILEO (EU) constellations. This network of satellites is used to determine the precise location of the tracker and thus the horse. Speed calculation can then be deducted by tracking the horse's position on the track at regular intervals and the change in location over time.

Page 37: Add here all terminology you are going to be using in your work: cardiac elasticity HRR, HR peak, stride length, speed etc... Add a brief table.
Table 1 added on page 41.

How those parameters compare to others used in different studies?
Comparisons were drawn between throughout the different chapters of the thesis.

Page 40: Include a section on statistical models used to compare large data sets; pros and cons, short paragraph.
Subsection added on page 40-41:
When comparing large datasets of animals retrospectively, various statistical models may be employed depending on the specific research questions and characteristics of the data. Commonly used models are briefly described below along with their main advantages and disadvantages.

Cox Proportional Hazards Regression, or Cox regression, is a statistical method used in survival analysis to examine the association between covariates and the time until a specific event (e.g., death or disease) occurs. It allows for the incorporation of time-dependent covariates and censoring. However, interpretation can be challenging for individuals not familiar with survival analysis concepts. Assumptions such as proportional hazards need to be assessed and, if violated, may lead to biased results.

Logistic regression is a statistical model used to analyse the association between a binary outcome (e.g., alive/dead, success/failure) or categorical (e.g., disease/no disease) and one or more predictor variables, estimating the probability of the outcome. It provides interpretable odds ratios that quantify the association between predictor variables and the outcome. However, it may not capture time-to-event data directly and may not be suitable if the outcome is not binary or categorical.

Generalised Linear Models is a flexible class of statistical models that extends linear regression to handle non-normally distributed response variables by specifying a link function and a probability distribution. Such models can accommodate various types of outcome variables (e.g., continuous, binary, count data) by specifying different distributions and link functions (e.g. Poisson and Logarithm for count data). However, model selection and interpretation can be complex, particularly with large datasets and many potential predictors.

Propensity Score Matching can be used to reduce selection bias in observational studies by matching treated and control subjects based on their propensity scores (i.e., the probability of receiving the treatment given covariates). It can be useful for comparing outcomes between treated and control groups in retrospective studies. However, it requires careful selection and matching of covariates, and the validity of the results depends on the adequacy of the propensity score model and the quality of the matching.

Time series analysis techniques can be used to analyse longitudinal data collected over time, such as repeated measurements on the same animals. However, such models may require a large number of observations to accurately estimate parameters, and they assume stationarity and independence of observations, which may not always hold in animal data.

Machine learning techniques like random forests, support vector machines, or deep learning can handle complex, high-dimensional data and capture nonlinear relationships between predictors and outcomes. However, such models may require a large amount of data for training and validation, and they can be computationally intensive. Interpretability may also be a concern with some machine learning models.

The choice of statistical model will depend on the specific research question, the type of outcome variable, the structure of the data (e.g., longitudinal, time-to-event), and considerations regarding model assumptions, interpretability, and computational resources. It is necessary to assess the performance of different models using measures such as predictive accuracy, goodness-of-fit statistics, and interpretability of results. Additionally, sensitivity analyses and validation techniques can help assess the robustness and generalisability of the findings.

Page 41: You have to clearly state that your work is an epidemiological retrospective study?
We did not believe this was the case, as per definition: epidemiology studies are research investigations designed to understand the patterns, causes, and effects of health and disease conditions within populations. These studies typically involve the collection and analysis of data from a group of individuals to identify associations, risk factors, and trends related to a particular health outcome or disease.
Whereas our study was an observational retrospective study, looking at factors that could affect racehorse performance rather than looking at what environmental variables that could affect their health. However, if you insist we can call it that.

What is data science?
Definition added: Data science involves extracting insights from data using a combination of statistical analysis, machine learning, and visualization techniques to ultimately inform decision-making.

Page 44: What do you mean by immersive study?
Added the word ‘comprehensive’ in the phrase to enhance understanding.
An immersive study refers to a comprehensive and deeply involved investigation into a specific research topic or area, often involving extensive hands-on experimentation, data collection, and analysis. It typically entails a thorough exploration of the subject matter, utilizing various research methodologies and techniques to gain a comprehensive understanding of the research question under investigation.

Are you comparing age vs HR?
Yes, this has indeed been done in the study, see Figure 4.

Page 46: Is this a correct reference?
No, reference removed.

Please clearly state number of horses included in each country and how many sessions per horse.
This is already stated on page 49-51.

This sentence is difficult to read, what do the numbers mean? Rephrase please.
Page 49, sentence rephrased to: The horses were 2 – 7 years of age. However, date of birth and thus racing age were only partially available (two; n=56, three; n=28, four; n=19, five; n=11, six; n=6, and seven-year-olds; n=3).

Page 47: How do the trainers detect the different horses? Clarify or add to limitations of the study.
Clarified on Page 49: Training usually occurred before daylight, with the trainers monitoring the horses from a distance, often from a watchtower, using binoculars. After training, the recorded Equimetre sessions were downloaded onto a laptop and reviewed by the trainer for each individual horse. There was no need to 'detect' horses as the training data wis made available after the exercise session.

Page 48: Please explain well how trainers did the categorisation of paces.
Categorisation of pace already explained on Page 49: 'With target timings to cover set distances speed could be known at any point along the track during each session, but to an extent was dependant on the jockey following the instructions of the trainer.' Added, 'to the their discretion'.

Please explain clearly what was considered hard, soft canter, etc in your study--difficult to compare with previous studies.
The speeds (i.e time taken to cover best 200m, max speed in km/h) of each pace are presented in Table 4 page 57 and varied by country. They were simply terms used to describe relatively slow-medium-fast canter// slow-medium-fast gallop. Those were the terms used by the Australian racehorse trainer and performance analysts and thus we also adopted them.

Any true validation of speed/distance data?
Page 42, paragraph added.
The device constructor reported a high accuracy of < 2.5 meters for distance in 50% of the trials and a high accuracy of 1.5meters @CEP (Circular Error Probable), open sky conditions during 24 hours. Thus, there is a high probability that the actual location or measurement will fall within a circle with a radius of 1.5 meters around the target or intended position, indicating the precision or reliability of the Equimetre system in determining locations or measurements.

Page 49: Be careful with the use of random. Please rephrase.
Not rephrased because any missing data was truly missing at random. That is what REML Model assumes in how the stats model is built.

Page 51: amend please, units incorrect.
Altered to metres.

Page 52: Discuss CV allowance, any diff between early and late collection during racing/training session? add explanation on discussion.

For the majority of parameters, the CV was very low (<5%). Indeed, it is relatively high for HR measures in that individual horse and can be affected by multiple factors. This was highlighted in the Discussion.

Page 55: Provide definition of paces.

The data-driven paces are already described in Table 4, Page 57.

As in any of the legends of your figures, re-phrase the text very difficult to read and follow throughout the Thesis.

The legends describe the data recorded accurately and the description that is written has been accepted by peer-reviewers and considered as acceptable for publication. Why should we make it more reader-friendly for a thesis?

Page 57: Define this term in your literature review and use it consistently through your work. remove cardiac elasticity.

Definition added to Table 1: the ability of the horse's cardiovascular system to quickly adjust heart rate during recovery from exercise, with higher values indicating a greater ability to return heart rate to baseline. The term cardiac elasticity/flexibility is akin to metabolic elasticity/flexibility and is just a term used to describe when an organ (e.g. heart) or system (e.g. metabolism) can readily transition between states; a lack of elasticity/flexibility implies propensity to various diseases.

Page 63: Please revise this relationship (sigmoidal) and amend as appropriate.

Page 65: sentence rephrased: However, when measured in situ with a rider aboard, and with racehorses taken to their cardiovascular peak (e.g., hard gallop), then heart rate measures were linear and tailed off at higher speeds. However, nearly all stimulus-response relationships in physiology are indeed sigmoid – much as we show for the heart rate response to exercise.

Page 64: Were horses already fit? so diff is ought to be minimal?

Yes, this is possible. Phrase added: Another explanation could lie in the fact that, for the recorded sessions, horses may have already been fit, which could explain the minimal difference.

Definition of 'Fitness' in horses: Reference?

Page 67, definition added: fitness may be defined as the ability to perform prolonged, continuous work - physical work capacity. Reference added. McMiken, D.F., 1983. An energetic basis of equine performance. Equine veterinary journal, 15(2), pp.123-133.

Page 68: Describe number of horses versus number of races.

This is already described in the abstract, Page 71: A total of 530 Thoroughbreds, varying in age (2-7 years old) and sex (including geldings), from one racing yard in Australia, completed a total of 3,568 exercise sessions, monitored by a single trainer, on varying track surfaces (sand, turf, or fibre). Finally, analysis of 52,464 race results indicated a similar chance of a top-three placing for male and female jockeys.

Page 72: Any differences between riders, age, experience etc?

This information was not available to us, hence sentence added: 'Other information such as rider's age and weight was not available.' We did not have ethics permission to get such 'personal 'information and the training team in Australia did not really want to ask weights of all participants.

Page 74: Explain what this parameter means, please % of race wins.

Definition added: '(i.e number of races won divided by total number of races, expressed as a percentage)

Page 76: What happened in 2020 versus 2021?
Added in Discussion part page 87: 'Additionally, the greater number of training sessions in 2021, may be explained by a possible increase in the number of Equimetre trackers at hand.'
Again, clarify legend by graph a) ...., b) ....
This is already specified: Figure. **a,b,e,f,g:** Data are predicted mean ± S.E.M. as recorded by 'Equimetre' in Australia (n=275 different racehorses, n=1,124 different training sessions) and in France (n=234 different racehorses, n=6,016 different training sessions). **c,d:** violin plots are all data with median and IQR indicated by solid and dashed lines, respectively. Training intensity (soft/med/hard canter; soft/med/hard gallop) was calculated from both cohorts based upon the fastest furlong (200m interval) for each session, excluding 'jumpout', for which there was no equivalent in France. Predicted means were calculated with training intensity fitted first, adjusted for sex of the horse and with track surface fitted last in the model. HorseID, racetrack and racehorse trainer were included as random effects in the model. All data analyses were conducted using Genstat

Page 79: Check legends.
This is already specified: Figure. **a,d:** Values are predicted mean ± S.E.M. for continuous data recorded by 'Equimetre' in a cohort of racehorses in Australia (n=130 different racehorses, n=1,754 different training sessions). Data were available throughout the year. Training intensity (soft/med/hard canter; soft/med/hard gallop) was derived from calculating sextiles of the fastest furlong (200m interval) for the overall dataset. The statistical model generated predicted means (± S.E.M.) with the pre-specified interaction, training intensity × rider sex (or training session**, e,f:** fitted last after inclusion of rider registration status and track surface as fixed effects, HorseID and rider name as random effects, since both horses and riders completed multiple sessions. All data analyses were conducted using Genstat v20 (VSNi, UK) and graphs produced using Graphpad Prism v9.0 (La Jolla, USA).

Page 80: Expand and rephrase " appeared influence. avoid appeared and trend in your discussion.
Page 82: Phrase altered to 'Over the period of greatest rate of recovery (i.e., to 15mins), the delta HR ([peakHR minus HR at 15mins]/15) showed differences based on the sex of the rider on turf, but not sand'.

Page 87: On balance, discernible? Simplify this sentence.
Page 89, rephrased to: 'Hence, based on our analysis, we conclude that the sex of the rider does not significantly influence racehorse performance in competitive races.'

Page 89: Since speed is determined by trainer, can this have any effect on gender?
This was indeed highlighted in our study: female jockeys did the vast majority of canters, whereas male jockeys did the majority of gallops (see Figure 5).

Page 90: Explain how, check the table in results as is confusing. explain % of wins parameter.
Table annotated for better understanding. See explanation for page 96.

Page 94: Explain your variables here: stride, peak stride, frequency of stride.
These are already described in the Table 1.

What for do you use your pedigree data?
Peak stride length and frequency.

Lifetime starts or starts over a lifetime? Please clarify.
Here, only race starts that were extracted and made available to us.

Page 96: Is your data really supporting this statement?

Yes, there were fewer stayers in our study (i.e 149 races out of a total of 3269 races; of which 26% were won compared to 1671 sprinter races, of which 14% won).

Is there class x distance interaction? Clarify in your discussion at least.
This was of interest however, we did not have enough data relating to the higher class-races to be able to explore the interaction, which would further increase the degrees of freedom.

Page 97: Explain this table better, numbers are confusing.
Table labels added.
What is top 3/ top 5?
Top3 (i.e ranking among the top three finishers) or top5 (i.e ranking among the top five finishers).

Ought to be explained here regression graph might have been better way of showing some of this data?
The table is clear to us and was the preferred choice. E.g of a total of 3269 races, 1671 were sprinter races, where 14% of those races were won.

Page 102: Avoid ‘trend’.
Removed.

Page 105: Is your data inconclusive or supportive?
Phrase changed to: ‘Whilst locomotory parameters have moderate heritability, locomotory parameters in individual racehorses do not evolve significantly during the course of a race season.’

Page 111: 3200?
Yes.

Page 114: Was class of horse a factor?
Yes. This is already described on Page 119: ‘Using this method, age, sex of horse, track condition, race-class were all included in any outcome.’

Page 115: Again, is random correct here? Who did the allocation?
Phrase added on Page 117: ‘to the discretion of the trainer’.

Page 117: Explain how this accounts for non-independent data from the same horse?
Page 120, sentence added: To account for multiple training sessions in the same horse then training session was coded (1 – 25) by date and analyses were blocked by HorseID.

Page 118: Experimental or comparison wise?
Statistical significance was accepted between comparisons.

How did you add inclination to the surface?
We did not have this information.

Page 120: What do groups represent? Add the different age groups in the legend.
Given that there was no real overall effect of age, we have chosen not to label the different age groups as it would make the graph too cumbersome.

Is the difference between 125 and 150 psychologically important?
Uncertain what you mean by this.

Unclear and unprecise, which difference?
In heart rate.

Page 121: Is fastest gallop 210 or 180, check both Fig 13 and 14.
We took the average at fast gallop, since as that time point both groups were identical as they also were for HRpeak. Therefore, at HR at1min it didn't matter which one we chose as the relative difference would be the same.

No error bars in any of your data points?
For many, the error bars are too small to be seen on the graph axis that shows the effect.

Please add number after figure in all your legends.
The legends describe the data recorded accurately and the description that is written has been accepted by peer-reviewers and considered as acceptable for publication. Unsure as to why we should make it more reader-friendly for a thesis.

What model did you use for analysis?
This is consistently stated in the legends of all graphs and tables.

Page 122: There is no gender in animals, sex.
Altered to sex.

Statistics?
Phrase altered to 'When accounting for racehorse age, sex, days-in-training and number of training sessions completed, then horses recovered better during hotter temperatures'.

How this links to your figures? same data or different?
This is clearly mentioned in the sentence with ref to the figure 16a.

Page 123: Mean difference? How predicted?
The predicted mean is weighted by the other factors (i.e age, sex, number of sessions) of the statistical models.

Page 124: What do you mean here?
Rephrased to: 'Overall, for stride length, a 0.17m difference existed between slow and fast horses.'

Is speed= stride length x frequency?
Yes.

Similar to before Table 11: clarify.
Page 127: Table annotated to 'slow-fast speed categories' for improved understanding.

Page 125: No error bars in your data sets.
Please discuss your significant difference data.
For many, the error bars are too small to be seen on the graph axis that shows the effect.

Page 129: Please explain % chance versus odds ratio in your discussion.
Explanation added on Page 136: To explore the correlation of physiological and horse-related measures with performance outcomes, two measures were employed in the statistical tests: percentage chance and odds ratio. Percentage chance, also known as probability, represents the likelihood of an event occurring as a percentage, calculated by dividing the number of favourable outcomes by the total number of possible outcomes. On the other hand, odds ratio compares the likelihood of an event occurring to the likelihood of it not occurring, expressed as a ratio of two odds. The odds of an event occurring are determined by dividing the probability of the event by the probability of it not occurring. While percentage chance provides

a direct measure of likelihood as a percentage, odds ratio offers a comparison between the chances of the event happening versus not happening.

Page 130: Explain this better.
Phrased altered to: In addition, horses recovered better in relatively hot, but not relatively dry, weather conditions.

Support or reject, amend this sentence.
Sentence altered to: This study provides evidence to support our primary hypothesis; training data can be used to predict which racehorses will a) go on to participate in a group race and b) be relatively more successful. Nevertheless, the absolute difference in practice is small (~4-6 beats/min; ~1-1.5sec/furlong) such that monitoring may not be of practicable use.

Page 131: No contractions in formal English.
Contraction removed.

Page 132: Different weather conditions? How this relate to Brazil Olympics?
Page 136, two relevant paragraphs added in this section.
Since the World Equestrian Games in Stockholm (1990) and the Olympic Games in Barcelona (1992), heat stress in horses has received significant media attention along with a surge in studies investigating how to preserve equestrian competitions in challenging environmental conditions (Jeffcott and Kohn, 1999, Marlin et al., 1996, Marlin et al., 1995, Kohn et al., 1995, McCutcheon et al., 1995, Lindinger et al., 1995).

Work by Munsters et al. (2024) evidenced that training for 14 days in a heated indoor arena contributed to the reduction of thermal strain on elite sport horses, facilitating competing in hot weather. The study also highlighted that high-level horses can acclimate to heat while remaining in training, a crucial factor for their participation in events such as the Olympics.

Page 134: Again, use throughout of trend and appear, use your data and stats significance.
Phrase changed to: 'Overall, stayers had a greater chance of winning relative to other categories of racehorse (sprinters and milers), as previously published.'

Page 135: Is this a prediction? detection?
Rephrased to: In conclusion, this study demonstrates that in-training racehorse physiological data can be used to identify which horses are more likely to win or not.

Page 136: Your discussion must clearly focus on your results and how those link to known literature as well as future work that can be developed from it. Please in detail revision of this chapter.
This discussion already discusses findings and existing work. I felt like I had discussed my results, as highlighted in the examples below:
- on page 142 for **Study 1**: 'Racehorses increased speed predominantly by an increment in stride length, then frequency, as also demonstrated in a study by Ratzlaff et al (1985). In both countries, speed was lower on sand, stride values (stride length and frequency) were reduced as previously reported (Rogers and Firth, 2004), but horses appeared to recover better when compared to turf. Peak heart rate and recovery increased with training intensity, with some differences observed between countries. We assumed this may be due to differences in training strategies between countries, as outlined in previous work conducted in Australasia, USA and the UK (Firth and Rogers, 2005, Morrice-West et al., 2020, Pagan et al., 2017, Verheyen et al., 2006b, Hodgson, 2014). Older horses recording lower heart rate at 15mins after training, which may be explained by the effect of age and/or improved fitness (Ohmura and Jones, 2017, Poole and Erickson, 2014, Yamanobe et al., 1993).'

-on page 143 for **Study 2**: 'Sex of the rider did not influence any aspect of racehorse speed, stride length, heart rate and peak heart rate at any training intensity. This is interesting considering other aspects have shown to influence speed or stride such as riding style (Pfau et al., 2009), weight (Gunnarsson et al., 2017) and rider's experience (Kapaun et al., 1998). However, racehorse heart rate recovery was influenced by sex of the rider when looking at the extremes of the reversed usual training intensity, that is: heart rate after galloping on sand was significantly lower with male riders. This observation draws attention to the nature of the horse-rider partnership dimension. Studies have explored the interactions between the cardiovascular responses of rider and horse in the context of fear/distress (Merkies et al., 2014), novel stimuli (Munsters et al., 2012) or rider (Schmidt et al., 2010)'

-on page 144 for **Study 3**: 'Our study found that, in training, sprinters had a shorter stride of higher frequency and covered consecutive furlongs faster than stayers. However, we were unable to account for the effect of warm up (Tranquille et al., 2017) and fatigue (Johnston et al., 1999), aspects that have previously been reported to influence stride characteristics. Relatively short or longer stride did not predict race success in sprint versus staying races, respectively. This is surprising as muscle strength, speed and endurance are previously identified traits that contribute to superior performance in Thoroughbreds at various race distances (Kay and Vamplew, 2012).'

-on page 144 for **Study 4**: 'Of all aspects tested, colts, stayers and fast finish speed were predictive of race performance and aligned with previously published work (More, 1999, Schrurs et al., 2022a, Kay and Vamplew, 2012). Opposingly, heart rate recovery and speed at the 600m mark did not show signals of performance prediction. The available literature on post exercise HR for performance prediction is somewhat contradictory as it includes either positive or negative corelations (Wilson and McGowan, 2019, Evans et al., 1993). It would have been interesting to account for the influence of the trainer, as it has been evoked to significantly associate with racing performance (Ely et al., 2010).'

Page 138: Is this sex or gender?
Sex.

Page 139: Please check your English here, avoid question marks in discussion.
Rephrased.

Page 140: How do you achieve your hypothesis?
Sentence removed.

Page 141: Add all limitations discussed throughout the corrections.
Page 146, added the below points:
-At the time of writing, we could only make use of what was available.
-To some extent, we could not account for horses being already 'fit' when evaluating various aspects of physiological responses to exercise.
-Additionally, speed categorisation, regardless of taking a data-driven approach, was to some extent dependent on the trainers chosen instructions.

Page 142: Future work: Use of treadmill to evaluate horse parameters? better or worse than in -field?
Page 146, paragraph added:
In comparison to treadmills, wearable technologies now allow for real-time monitoring of racehorse performance in their natural environment, with little to no interference with their usual work routine. However, certain challenges persist for with field sampling, such as blood or lactate measurement, due to their vivid demeanour of racehorses. Therefore, such measurements may still be better suited for treadmill conditions.

Page 142: How your data/device can improve horse training and horse welfare?
Put more emphasis in how your work could lead to other.

Page 147, paragraph added:

In essence, the integration of objective fitness trackers shapes a promising new era for improved racehorse health and welfare on the track. With 'live' data on the track, such innovative devices offer real-time monitoring capabilities allowing trainers and veterinarians to gather precise data on various aspects of equine physiology and performance. By tracking parameters such as heart rate, ECG and stride, fitness trackers provide invaluable feedback that can inform more targeted training regimens and identify early signs of stress, injury or fatigue.

Page 143: Conclusion: re-write based on your most important outcomes/findings.

Page 148, paragraph added.

Our work highlighted that track surface influenced racehorse speed, stride and heart rate recovery (i.e lower values on sand compared to turf). Older horses presented lower heart rate values. Overall, there were marked differences in training strategies between countries, with Australia conducting shorter intervals of distance at faster speeds, compared to France, where a more conventional training approach was adopted.

The sex of jockey did not appear to influence speed, stride length nor heart rate. Both male and female jockeys presented a similar chance for top 3 ranking and win percentage.

When comparing locomotory racehorse profiles, then sprinters presented a shorter stride but of increased frequency, being on average faster than stayers. Short and long stride did not predict racing success. With all aspects considered, then stayers had greater chances of success, with peak stride length/frequency being moderately heritable.

Lastly, analysis of cardiovascular fitness and stride acceleration in race-pace workouts for the prediction of performance outlined that aspects such as colts, stayers and fast finish speed were predictive as opposed to heart rate recovery and speed at the 600m distance mark.