Chapter 1

General Introduction
1.1 Professional road racing

The first bicycles were introduced in the late 19th century in Europe. Directly after the introduction people wanted to see how fast they could cycle and started racing against each other. Soon the first bicycle races were born. The first recorded official bicycle race was organized on May 31 in the year 1868 at the Parc de Saint-Cloud, in Paris. A 1.2 km race was won by an Englishman named James Moore and this first race was promoted by the Michaux company. In those early days the races were short, because of the poor user-comfort of the bicycles and the poor materials the bicycles were made of. However, with more improvements and innovations to bicycles, the races became longer and longer. The first recorded official ‘long’ road race was a 199 km race from Milan to Turin in 1876, a race that still exists. This was followed by Liege-Bastogne-Liege in 1892 and Paris-Roubaix in 1896. Both of these races still belong to the most prestigious races in modern road cycling. In 1896, bicycle racing was introduced during the first modern Olympic games. In 1903 the first Tour de France was organized, consisting of six separate stages, covering 1500 km. The race started in Paris and moved to Lyon, Marseille, Bordeaux and Nantes before finishing in Paris. The first winner was Maurice Garin. Nowadays, the Tour de France is one of the largest sports events in the world, with over 12 million spectators, visiting one of the 21-stages live while the race is broadcasted in 190 countries.¹

Modern competitive road cycling is one of the sports with the highest physiological demands. A professional male cyclist will cycle in training and races between 30,000 and 35,000 km per season.² In 2011 the Union Cycliste Internationale (UCI) introduced the World Tour, a combination of the most prestigious 34 – 40 (depending on the year) one day and multiple-stage races. Those races are spread over four continents (Asia, Australia, Europe and North America) and interest is growing in the other continents (Africa and South America). There are 18 World Tour teams which do not depend any more on invitations for important races but are obligated to start in every World Tour race. The new World Tour results in an almost all year around competitive season. A professional male cyclist generally races from January to the beginning of October, with the off-season starting in October or November, the cyclist is back in training at the latest in the beginning of December.
Road racing is a complex sport. Although only the winner of the race can participate in the podium ceremony, road cycling can be seen as a team sport. Drafting behind domestiques (support riders) is highly important and can save a team leader up to 40% of energy compared to riding alone.\textsuperscript{3,4} Further, domestiques are used for positioning in the peloton at the key moments in the race and help with spare materials, nutrition and hydration so the team leader can save energy.

A professional male cyclist generally has between 60 and 100 competition days, depending on their age, experience and speciality. A neo-prof (first year professional) will have a lower number of competition days compared to more experienced riders. Further, domestiques will race more compared to leaders, because the leaders are optimally prepared for specific races where they hope to reach their peak performance level at the right time. Those 60 to 100 Race-days include 1-day races (from 180 km up to 300 km), multiple-stage races (5-10 race-days) and Grand Tours (GTs) (21 race-days).

Over the last decade, women's cycling gained more attention from the public, the UCI and sponsors, which has resulted in their own Women's World Tour in 2016. Although published data on training and race characteristics are limited, our own observations have shown that female World Tour cyclists will typically cover between 13.000 to 18.000 km in training and competition. With the new World Tour, women’s races are organized in Europe and Asia, with upcoming races in Australia and North America.

Professional women cyclists race up to 65 race days in competition. Lacking an under-23 category, junior female cyclists already start racing with the senior women from the age of 18, in contrast with male cyclists who have an under-23 category. Race programs are based on age and experience and thus lacking an under-23 category results in large differences in the amount of race days of female cyclists. Women’s races have a maximum of 160 km for the Olympic road race and Word Championships, the races on the World Tour are limited to a maximum of 140 km. In contrast to men’s cycling, there are no GTs in female cycling and the longest multiple-stage race is the Giro D’Italia Internazionale Femminile which consist of 10 days of racing without a rest day.
1.2 The training process

Training is a process whereby athletes are systematically exposed to stimuli with the goal of inducing adaptation in the performance determining factors, in order to improve skill and performances. Within this training process, overcoming training and competition stresses promotes will power, self-confidence and tolerance for higher training and competition loads. The training session induces a physiological response, and it is this response rather than the exercise task itself that provides the stimulus for adaptation. The training process can be divided into different periods from long term training spans (10-15 years) to shorter periods like a season or even shorter phases like one week or even one day. Optimal performance strategy centers around the problem of how to design a training program within an annual plan that 1) maximizes performance potential at a known date and 2) minimizes the risk of fatigue, overtraining and injures during the period of training leading up to the target date. Periodization is a process of planning that enables the utilization of correct loads and adequate recovery periods in order to avoid excessive fatigue. Coaches develop a template for the athlete, which is preferably not rigid but a dynamic framework, which can be adjusted for specific situations. Periodization provides the structure for controlling the stress and recovery that is essential for positive training response. Coaches can manipulate the Frequency, Intensity, Time (duration) and Type (FITT principle) of the training sessions to achieve different training goals and build a balanced training program (Figure 1.1).

1.3 Training load and Intensity measures in professional cycling

Training programs are based on a certain amount of Training Load (TL) or training dose. Performance improves with gradually increasing TL, although too much TL is associated with the overtraining syndrome and/or negative adaptions to exercise. In order to avoid under- and over-training, coaches monitor TL and adjust the load throughout the training program to either increase or decrease fatigue depending on the phase of training. Each indi-
individual athlete has a unique reaction to a certain TL. This unique reaction depends on the individual's capacity, which can be seen as a dynamic process. This means, for example, that the capacity of an athlete can change during a training session, day or week. Further, the capacity of an athlete is influenced by physical as well as the psychological status of the athlete. Therefore, monitoring TL and the influence of TL on the individual response is highly important. With this knowledge, it is possible to develop a training program for the individual athlete which is tailored to the athlete’s individual response.

TL can be distinguished in external and internal TL. External TL is defined as workload done independently of the athlete’s internal characteristics. Therefore, two athletes may undertake an identical training when expressed in external loads but experience quite different internal load, depending on their individual characteristics, training status, psychological status, health, nutrition, environment and genetics. Examples of external

![Figure 1.1: Theoretical framework between training goals, external and internal training load and the training outcome. Adapted from Impellizzeri.](image)
load in cycling are distance (km), duration (min) and mechanical energy spent (kJ spent). Internal workloads quantify the physical loading experienced by the athlete. An internal load is a subjective index of the effort and internal loads are the response of an athlete on the external load.\textsuperscript{14} Examples of internal load in cycling are Heart Rate (HR), HR derivates (luTRIMP, eTRIMP) and Rating of Perceived Exertion (RPE). It is suggested that a combination of external and internal load measurements are important for training monitoring and performance prediction.\textsuperscript{15} However, the focus should mainly be on monitoring internal load rather than external load, as internal load ultimately determines the functional outcome of training (i.e. positive, or negative, adaptation) (Figure 1.1).\textsuperscript{9} Furthermore, Sanders suggested that load values that integrate individual physiological characteristics have the strongest dose-response relationship in professional cyclists.\textsuperscript{16} There are primarily two ways to analyze TL, absolute and relative. Absolute load is the sum of all training sessions in a certain period, or training period, and is mostly described per day, per week (acute load) or per 4 weeks (chronic load).\textsuperscript{14} Relative load is the load expressed as a percentage or ratio to an amount of historical load (i.e. load in the last month or week). Absolute and relative loads can be measured from both internal and external load measures.

Several internal and external load measures have been reported in professional cycling. Most load measurements have somehow integrated the intensity and duration of the exercise in their value. Highly used in professional sports as internal load is session RPE (sRPE) introduced by Foster in 1995.\textsuperscript{10,17} This is a multiplication of the RPE\textsuperscript{18} by the duration. The sRPE is popular in sports because it is noninvasive, inexpensive and very easy to use. HR is also frequently used to determine internal TL. Similar to RPE, HR is noninvasive and inexpensive and therefore widely used. There are different methods to calculate TL based on HR. Banister et al.\textsuperscript{19} introduced the concept of the training impulse or TRIMP, a TL measure based on the intensity of the exercise as calculated by the product of the average heart rate reserve (%HRreserve) and the duration of the exercise.\textsuperscript{19} Today, several variations on TRIMP are described in literature. In 1993, Edwards proposed a simplified measurement of TRIMP based on 5 defined HR zones (eTRIMP),\textsuperscript{20} where the duration spent in each zone is multiplied by a weighting factor from 1 to 5, and then summated to provide a total eTRIMP score.\textsuperscript{20} Highly used in professional cycling is a
TRIMP score proposed by Lucia et al.\textsuperscript{21,22} which is based on the three HR zones, around the two ventilatory thresholds. Zone 1 is below the Ventilatory Threshold (VT), zone 2 between the VT and the Respiratory Compensation Point (RCP) and zone 3 above the RCP. Comparable with eTRIMP, time spent in each zone is multiplied with an arbitrary factor from 1 to 3 and then summated to provide one total score of Lucia’s TRIMP (luTRIMP). In the past, cycling training programs were mostly based on HR, RPE or external load measures such as duration and distance. However, with the introduction of the power meter mounted on the bicycle, multiple external load measures came in use, such as the mean power output and kJ spent. In 2003, Allen and Coggan\textsuperscript{23} introduced a new power-based TL called Training Stress Score or TSS and this is now widely used in professional cycling, mainly because the most popular training software on the market (i.e. Trainingpeaks) is based on TSS. TSS is based on the Functional Threshold Power (FTP) of a cyclist in which a ride of 1 hour at FTP is equal to 100 TSS. FTP represents thus the nominal maximal power output a rider can sustain for 1 hour.

Training intensity is one of the four principles of the FITT principle and coaches can adjust the training intensity to influence training goals and the training outcome (Figure 1.1). Quantifying training intensity is more complicated compared to quantifying TL. Mostly in sports the training intensity is a scale based on ranges of HR relative to the maximum HR or to blood lactate concentration.\textsuperscript{24} Training intensity zones are mostly divided into five intensity zones or three intensity zone. The intensity zones are somewhat arbitrary and are criticized because this approach lacks the individual variations between HR and blood lactate concentrations.\textsuperscript{25} However, in the literature, also individually determined intensity zones are studies, based on three intensity zones anchored around the first and second ventilatory thresholds.\textsuperscript{21,22,25-27} In professional cycling, intensity zones are based on HR,\textsuperscript{21,22,26,28} RPE\textsuperscript{26,27} and PO.\textsuperscript{27} Sanders et al.\textsuperscript{27} compared the individually determined intensity zones based on HR, RPE and PO in professional cyclists and concluded that intensity distribution based on sRPE differs compared to the intensity distribution based on HR or PO. Allen and Coggan\textsuperscript{23} introduced a seven intensity zone distribution based on a percentage of FTP. These power-based zones are somewhat arbitrary chosen based on RPE and HR. Although lacking a scientific basis, this approach is highly used in professional cycling because it is implemented in
the most popular analyzing software (i.e. TrainingPeaks).

The intensity distribution of a training session or training program is highly important. Coaches manipulate the distribution of training intensity to either increase load and/or to differ in training goals. Coaches have several options to structure the intensity distribution and different approaches are described in the literature (e.g. threshold-based approach, polarized approach and pyramidal approach). However, research indicate that a polarized training approach is the most effective in elite endurance athletes in comparison with other training approaches. A polarized training approach is characterized by an intensity distribution that consist of ~80% of the training sessions performed at zone 1 and the remaining ~20% at zone 3, with little or none in zone 2. It is important to note that some TL principles described above have integrated intensity by multiply training session time with physiological or perceptual measure of intensity (i.e. luTRIMP, eTRIMP, sRPE and TSS). However, in this thesis, training intensity is described as the distribution in the intensity zones.

1.4 Load and intensity demands and performance in professional road cycling

In the last decades, technological advantages have made it possible to measure the load, intensity and performance demands of road cycling. With the introduction of the mobile heart rate monitor in 1977 by Polar Electro, professional athletes started to train based on their heart rate. However, in 1986 the mobile power meter was introduced by Uli Schoberer and in the recent decade this changed training and analysis of professional cycling from a heart rate-based to a power-based approach. As a result of those technological advantages, data of professional cyclists during racing and training are easy to collect and accessible for in depth analysis, with the result that multiple applied and descriptive studies have been published about the load, intensity and performance demands of professional cycling.

Road cycling is a unique sport wherein various race formats are embedded, which have their own physiological characteristics. Road cycling is a typically endurance sport and thus a cyclist’s maximal oxygen uptake
General Introduction

Chapter 1

The peak oxygen uptake (VO\(_{2\text{max}}\)) and the ability to sustain a high percentage of VO\(_{2\text{max}}\) is important.\(^{46,47}\) VO\(_{2\text{max}}\) values from elite male road cyclists have been reported between 5.0 to 5.5 L·min\(^{-1}\) or 70 to 80 mL·min\(^{-1}·kg\(^{-1}\)).\(^1,2\) Recently, physiological values of Tour de France winner Chris Froome were published and reported a VO\(_{2\text{max}}\) of 5.91 L·min\(^{-1}\) or 84 mL·min\(^{-1}·kg\(^{-1}\)) and a peak power output (PPO) of 525 W in a laboratory incremental cycling test.\(^{48}\) Reported VO\(_{2\text{max}}\) values in professional female cyclists are substantially lower compared to men, but averaging 65 mL·min\(^{-1}·kg\(^{-1}\)) are still very impressive.\(^{49}\) Although VO\(_{2\text{max}}\) is very important, Chicharro et al.\(^{50}\) showed that amateur and professional road cyclists had similar VO\(_{2\text{max}}\) values but that the distinction between the different performance levels was based on the VO\(_{2\text{max}}\) at the first and second ventilatory thresholds (the Performance VO\(_{2\text{max}}\)). The first and second ventilatory threshold in professional cyclist occur at a higher percentage of VO\(_{2\text{max}}\) (74% and 90% respectively) compared to the amateur cyclists (61% and 81%, respectively).\(^{51}\) Although road cycling is mostly an endurance based effort, the ability to tolerate constant workload around 90% VO\(_{2\text{max}}\) during prolonged periods of time\(^{21}\) and a high anaerobic capacity are necessary to compete for victory. The decisive moments in road cycling are mainly based on those capacities, because time trials (TT), closing a gap, break away from the peloton or winning a sprint are all efforts were endurance capacities alone are not enough.\(^{41,42,46,52}\) Cyclists that want to win road races need to have an exceptional combination of aerobic and anaerobic capacities.

Road races typically consist of 4 different types of races (e.g. flat, hilly, mountain and (team) TT. Although it is possible to distinguish even more subspecific between different types of stages e.g. the cobble classics, the Belgium Ardennes classics, team TTs and specific finishes for punchers which specializes in short but steep climbs. Most riders are specialized in one or two of these competition elements, because physiological and anthropometric demands differ between each specialty.\(^{2,53,54}\) In addition, the load and intensity demands of those specific types of races differ as well.\(^{35,38}\)

Flat stages are a key feature in road cycling, approximately one third of the stages in a multiple-stage race will finish in a bunch-sprint or a reduced bunch sprint (50-100 riders) and are specifically designed for sprinters.\(^2\) In addition, some 1-day races (i.e. classics) finish in a bunch or a reduced bunch sprint. Sprinters usually have a higher body weight and height compared to riders specialized in other disciplines.\(^{34}\) On flat terrain the primary force to
Chapter 1

overcome is the air resistance.\textsuperscript{2} Cyclists can draft the most of the time during flat stages, which reduces energetic requirements by as much as 40%.\textsuperscript{3,4} Thus, not surprisingly, load and intensity demands are lower in flat stages compared to the high mountain stages.\textsuperscript{35,38} However, high anaerobic capacities are needed for finishing in the top 5 or winning a sprint stage. This is illustrated by Menaspa et al.,\textsuperscript{41} who reported a peak power output of 17.4 W·kg\textsuperscript{-1} and an average power output of 14.2 W·kg\textsuperscript{-1} during a sprint with male sprinters. The power output of female sprinters is slightly lower, with a peak power output of 13.9 W·kg\textsuperscript{-1} and an average power output of 10.6 W·kg\textsuperscript{-1} during a sprint.\textsuperscript{43} Not only physical capacities are necessary to win sprint stages but also tactics, positioning and teamwork with domestiques are factors that increase the chances of winning a sprint stage.\textsuperscript{40}

High mountain stages, especially those with an uphill finish, are of the utmost importance in professional cycling. Most General Classifications (GCs) are decided during the high mountain stages or TTs. A high mountain stage contains on average 3814 meters of elevation gain\textsuperscript{35} and can include 3 to 5 mountain passes with an average 5 to 10\% gradient.\textsuperscript{2} In contrast to sprint stages, cyclists must mainly overcome the force of gravity during the high mountain stages and the ratio of power to body mass is extremely important.\textsuperscript{2} Both Vogt et al.\textsuperscript{38} and Sanders et al.\textsuperscript{35} studied the load demands of the different types of stages (flat vs mountain) in road cycling and reported that high mountain stages are significantly harder compared to flat stages. Professional cyclists specialized in high mountain stages are extremely lean. An average 7.8\% fat mass is reported for a Giro d’Italia team, including all different specialists (i.e. sprinters, climber and domestiques),\textsuperscript{55} thus cyclists specialized in climbing are probably even leaner. Measurements in the laboratory shows that climbers can sustain more than 6.0 W·kg\textsuperscript{-1} for longer time at maximum or near to maximal cycling.\textsuperscript{34,56} Further, it seems that professional women cyclists produce 5.3 W·kg\textsuperscript{-1} at the respiratory compensation point and 6.6 W·kg\textsuperscript{-1} at maximal intensities in a laboratory test.\textsuperscript{49} Although some laboratory values are known for professional cyclists, the exact performances of climbers during races are still unknown.

Another discipline in road cycling are TTs. TTs are usually crucial in the final outcome of multiple-stage races and GTs. Almost all multiple-stage races contain one or two TTs and some GTs contain even 3 TTs. As drafting
is not possible in an individual TT, the air resistance is the most important force to overcome. Cyclists seeking to win a TT must tolerate high intensities close to the ventilatory threshold (e.g. second ventilatory threshold) during the entire TT.\textsuperscript{2} Reported power output during a TT variates from 347 W to 380 W\textsuperscript{57} and TT specialists can produce power outputs higher than 400 W for longer time.\textsuperscript{54} Although high intensities during TTs, load values are lower during TTs compared to the sprint and high mountain stages, because of short durations.\textsuperscript{35}

A GT is the most demanding and prestigious discipline in professional road cycling. A typical GT contains of 21-race days with only 2 or 3 rest days. During a GT, between 3500 to 4000 km are covered in 85 to 95 hours with all the different stage types being part of those 21 days of racing.\textsuperscript{2} GTs have an extreme high physiological demand and multiple descriptive studies evaluate the load and intensity demands of GTs.\textsuperscript{22,35-37,39} Those high demands have a high physiological and psychological impact on the cyclists and research has shown that hormonal levels,\textsuperscript{33,50} sleep,\textsuperscript{58} mood,\textsuperscript{58} well being,\textsuperscript{58} heart rate\textsuperscript{22} and blood values\textsuperscript{59} are effected by those demands. In addition, Rodriguez-Marroyo\textsuperscript{60} showed that long term fatigue accumulated in professional cyclists over 21-days of racing resulted in a performance decline of $\sim$10\% and affected both maximal and submaximal endurance performance.

Understanding the load, intensity and performance demands in professional male and female cycling is necessary to compete on the highest level. Knowledge about the load, intensity and performance demands provide coaches and sport scientists the ability to build an optimal training program (Figure 1.1). The capacity to sustain high continuous loads and intensities during multiple-stage races are determinants which are necessary in professional cycling. A detailed training program should be based on the demands necessary in races.\textsuperscript{5} Although multiple studies evaluated the load and intensity demands in professional cycling, there are still multiple unknowns. Especially the load and intensity demand in women's professional cycling are less well documented and could differ from the demands in professional male cyclists. Further, in the recent years multiple studies have highlighted the demands of GTs and their different type of stages.\textsuperscript{21,22,34-36,38,39} However, the demands for a GC contender are still unknown.
1.5  **Training program**

Long term monitoring of training and performance of world-class athletes is rare and in professional cycling only a few cases are reported.\(^{45,61}\) Unfortunately, in one of these investigations the studied athlete (Lance Armstrong) admitted to using performance enhancing drugs and, therefore, the scientific validity of Coyle’s study cannot be confirmed. Pinot et al.\(^{45}\) monitored a top 10 GT finisher for 6 years and described the TL and performance characteristics. Within this study the importance of monitoring and planning the TL in developing an elite cyclist is described. He showed that the load increased progressively throughout the years and the cyclist progressively improved, with a gain of 14% in PO between the ages of 18 and 23 years.\(^{45}\) Sanders et al.\(^{16}\) focused their work on the dose-response relationship between TL and performance. They investigated the optimal much load for improving performance in elite cyclists and concluded that load measurement methods that integrate individual physiological characteristics have the strongest dose-response relationship. Although there are limited data describing the training characteristics of male professional cyclists, to the best of my knowledge there is no research providing a detailed quantification of the load and intensity demands of female professional cyclists. As argued above, it could be that the difference in race-program between men and women could lead to different demands in races. The capacity to sustain those demands is one of the determinants to be able to perform in races and therefore should be part of the training goals (Figure 1.1).

A successful training program requires a balance between overload and recovery, which results in positive adaptations and thus an increased performance.\(^{5,62}\) Banister et al.\(^{63}\) introduced in 1975 a model that explained this fine balance between load and recovery. When load and recovery are in balance, the result is a performance gain. However, when there is a disbalance between load and recovery it could lead to negative adaptations or a decrease in performance.\(^{63}\) Performance is related to the difference between the fitness response and the fatigue response. The fitness response is related to physiological adaptation to training and will result in an improvement.\(^{63}\) The fatigue response is a shorter response process and is related to the negative adaptation of training.\(^{63}\) Successful training must involve overload. However,
excessive overload in combination with inadequate recovery could result in negative adaptations. A single training session or overload training period could result in feelings of fatigue and a decrease in performance. However, combined with sufficient recovery, the applied load results in a positive adaptation and an increased performance. This contrasts to an increased load that is combined with insufficient recovery, which could lead to overreaching or in the long term to overtraining syndrome. A period with a high acute load (i.e. overload period) is also associated with an increased risk of injuries and illnesses in multiple sports. Periods with an acute high (external or internal) load could increase the risk of injury for up to 1 month. Further, Hulin et al. showed that the ratio between acute load and chronic load had a higher association with injury risk then acute load alone. In contrast to a high acute load, it seems that a high chronic load, provided that it is gradually built up, could protect against injuries. Comparable to injuries, similar relationships have been found between TL and illnesses, although the evidence is not as clear as the relationship between TL and injury. Elite athletes that can maintain training availability at 80% have a significantly greater chance of successfully achieving their key performance goals. Therefore, attention should be paid to the prevention of both injury and illness in order to maximize the likelihood of success in elite athletes.

1.6 Aim and outline of this thesis

Limited evidence is available about the load and intensity demands of races in professional cyclists, especially female cyclists. One of the determinants to perform in professional races is the ability to sustain high load and intensity demands. Knowledge about those demands could lead to new insights into the training goals in professional cycling. Furthermore, in other sports a rapid increase in load is associated with a higher risk of injuries and illnesses, while this is still unknown in a low impact but high-volume sport like professional cycling. Therefore, the aim of this thesis is to explore the framework (Figure 1.1) proposed by Impellizzeri et al. with data from professional cyclists. Understanding this framework in professional male and female cycling will lead to insights regarding the load and intensity demands and performance in professional cycling and the negative response that is induced
with a disbalance between load and recovery. This thesis will contain 6 studies reported in Chapters 2 to 7. A short outline of these studies is given below.

Chapter 2: Load, intensity and performance characteristics in Grand Tours.
Despite multiple studies describing the load and intensity demands of GTs, the load and intensity demands of competing for the victory in a GT are still unknown. This Chapter presents a case study of a GC contender describing the load and intensity demands in multiple GTs. Further, this Chapter highlights the cyclist’s performance during those GTs.

Chapter 3: Intensity and load characteristics of professional road cycling, differences between men and women.
Race regulations differ between men and women. Therefore, there are differences between the race formats. It is unknown if those different race formats also result in different load and intensity demands between women and men. The aim of this Chapter is to describe the load and intensity demands of professional races and to highlight the differences between male and female cyclists.

Chapter 4: Intensity and load characteristics of professional cyclists in training, differences between men and women.
Chapter 3 will reveal some interesting differences between the load and intensity demands of men’s and women’s races. Because demands in training are related to the demands in races (Figure 1.1), the results of Chapter 3 could lead up to different training programs for men and women professional cyclists. Although there are some limited reports about the training demands of professional male cyclists, the training demands of female cyclists are unknown. The aim of this Chapter is to describe the load and intensity characteristics in the training of professional cyclists and to highlight the differences between male and female cyclists.

Chapter 5: The relationship between various training load measures.
Various TL measures are used in professional cycling (i.e. sRPE, luTRIMP, TSS and kJ spent). The relationship between those TL measures is not well
explored. This Chapter will investigate the inter-relationship of sRPE, luTRIMP, TSS and kJ spent in training, road races and TT's. It is hypothesized that factors which influence TL are harder to control in competition, resulting in a weaker relationship between the various relationships.

Chapter 6: The influence of exercise intensity on various training load measures.
TSS is highly influenced by exercise intensity and the results from Chapter 5 will indicate differences between TSS and the other TL measures (kJ spent, sRPE, luTRIMP). Therefore, in this Chapter the influence of exercise intensity on various TL measures is studied. It was hypothesized that TSS reacts differently on exercise intensity compared with sRPE, kJ spent and luTRIMP.

Chapter 7: Training load in association with injuries and illnesses.
Multiple studies associate TL (ratios) with a higher risk of injuries and illnesses in different sports. Although frequently studied in other sports, there is no research published in which those associations in a high volume, low impact sport like professional cycling are studied. It is hypothesized that a high acute load and a high workload ratio have an increased risk of injuries and illnesses, while a high chronic load lowers the risk of injuries and illnesses.

Chapter 8: General discussion
This final Chapter will discuss the results presented in the different Chapters. In addition, future research directions are provided, together with practical implications.
Chapter 2

Case Report: Winning a Cycling Grand Tour, What does it take?

Teun van Erp • Marco Hoozemans • Carl Foster • Jos J. de Koning

*Medicine & Science in Sports & Exercise*

*(Accepted for publication)*
Abstract

Introduction: The aim of this study was to present the load, intensity and performance characteristics of a General Classification (GC) contender during multiple Grand Tours (GTs). This study also investigated which factors influence climbing performance.

Methods: Power Output (PO) data were collected from a GC contender from the Vuelta a España 2015, the Giro d’Italia 2017, the Giro d’Italia 2018 and the Tour de France 2018. Load (e.g. Training Stress Score and kJ spent) and intensity in 5 PO zones was quantified. One-way analysis of variance was used to identify differences between the GTs. Further, performance during the four GTs was quantified based on maximum mean power output (W·kg⁻¹) over different durations and by the relative PO (W·kg⁻¹) on the key mountains in the GTs. Stepwise multiple regression analysis was used to identify which factors influence relative PO on the key mountains.

Results: The athlete was highly successful during the four GTs in which he finished 6th, 1st, 2nd and 2nd in the GC. No significant differences were found between load and intensity characteristics between the four GTs with the exception that during the Giro d’Italia 2018 a significantly lower absolute time was spent in PO zone 5 (P=0.005) compared to the other three GTs. The average relative PO on the key mountains (n=33) was 5.9±0.6 W·kg⁻¹ and was negatively influenced by the duration of the climb and the total elevation gain before the key mountain, while the gradient of the mountain had a positive effect on relative PO.

Conclusions: The physiological load imposed on a GC contender did not differ between multiple GTs. Climbing performance was influenced by short-term fatigue induced by previous altitude meters in the stage and the duration and gradient of the mountain.

Keywords: general classification contender, professional, team leader, performance, power output.
2.1 Introduction

Elite road cycling is one of the most physically demanding sports in the world, especially the 3-week multi-stage races: the Giro d’Italia, the Tour de France and the Vuelta a España. A Grand Tour (GT) typically contains 21 race-days with only 2 or 3 rest days, covering between 3500-4000 km in 85-95 hours of competition.\(^2\) For elite cyclists, to participate in one of the three GTs is the highlight of the season and winning a GT is the highest achievement possible in professional cycling. A GT is a complex race with different competition elements, and therefore different kinds of stages (e.g. flat, semi-mountainous and mountain stages).\(^{35,39}\)

The introduction of the heart rate and Power Output (PO) monitors ensure that more information is obtained about the load and intensity demands of participating in a GT.\(^{22,35,39,57,60,69,70}\) In 2003, Lucia et al.\(^{22}\) showed with heart rate data of seven cyclists the load and intensity demands of two different GTs (the Tour de France and the Vuelta a España) and did not find any significant differences between the load and intensity despite the assumed different nature of both races. Vogt et al.\(^{39}\) and Sanders et al.\(^{35}\) discussed, on the basis of heart rate and PO, the load and intensity characteristics of different terrains (e.g. flat, semi-mountain and mountain stage) in GTs.\(^{35,39}\) In addition, Rodriguez-Marroyo (2017)\(^{60}\) showed that fatigue suppresses the heart rate during a GT and was able to correct the load and intensity demands of the Vuelta a España by adjusting heart rate thresholds based on pre- and post-laboratory tests. Although multiple studies report load and intensity characteristics of GTs, it is still unknown what the load and intensities are for a General Classification (GC) contender.

Winning a GT is the highest achievement possible for an elite cyclist and only accomplished by the best and most complete cyclists. With exceptions from (team) Time Trials (TTs), differences in the GCs are mostly made by performances uphill, which can be defined by the relative PO during that climb. Therefore, relative PO on the last mountain in a stage is very important for a GC contender. Pinot et al.\(^{45}\) presented the relative mean maximum PO over different durations (15-18000 seconds), the so called “Record Power Output”\(^{45}\) or “Maximum Power Profile”\(^{23}\) (MPP) of a GC contender achieved in training and competition. In addition, Sanders et al.\(^{35}\) presented the MPP
values of 9 riders during the Giro d’Italia. Both studies give a rare glimpse of the performances necessary to compete in a GT. However, PO values on the key mountains for a GC contender in a GT are still unknown.

Several factors can influence the performance of GC contender on the key mountains during a GT. Due to the very demanding nature of GTs, elite cyclists accumulate fatigue over the three weeks of racing and this may reflect in several physiological and psychological changes which affect performance.\textsuperscript{22,36,58,71} Recently, Rodriguez-Marroyo et al.\textsuperscript{60} showed that long term fatigue accumulated by professional cyclists over 21 days of racing resulted in a performance decline of \(~10\%\) and affected maximal and submaximal endurance performance. Further, it is expected that fatigue sustained during the stage (short term fatigue) has an influence on the performance at the end of a stage.\textsuperscript{72} Furthermore, mountain characteristics such as the length and the gradient of the mountain will have an influence on the performance and thus on relative PO.\textsuperscript{69,73} Lastly, environmental conditions such as heat could impair performances.\textsuperscript{74}

Although there is growing interest in the load, intensity and performance characteristics of elite cyclists competing in GTs,\textsuperscript{22,35,39,60} the reported values in the literature are based on the average of multiple cyclists. In GTs, teams select different types of riders, who have different roles throughout the race (e.g. sprinter, domestic and TT specialist) and thus don’t have to perform all 21 stages at their top level.\textsuperscript{53,54} This contrasts with a GC contender who must perform every day to prevent losses in time. Therefore, load and intensity demands for a GC contender could be higher than the reported values for the average team. However, there is evidence of “pacing” across the duration of GTs, whereby GC riders only ride at maximal effort during a few stages.\textsuperscript{75} In addition, to the best of the author’s knowledge no study reported performance characteristics of a GC contender or factors which could influence the performance of a GC contender during a GT. Therefore, our research question was three-fold: First, to present the load and intensity characteristics of a GC contender for multiple GTs. Second, to present the relative PO which is needed on the key mountains to compete for the victory in multiple GTs. Third, to examine which factors are associated with the relative PO on the key mountains.
2.2 Methods

Participant
The athlete is a professional cyclist (24 y, 69.9 kg, 185 cm) competing in the UCI World Tour Series and provided written informed consent for a detailed analysis of his PO data collected during the Vuelta a España 2015, the Giro d’Italia 2017, the Giro d’Italia 2018 and the Tour de France 2018, where the athlete competed for the GC victory. During the 7 years that the athlete was professional cyclist, he was highly successful and in 2017 and 2018 ranked as a top 10 cyclist in the world according to the World Tour ranking. The main specialty of the athlete are TTs, where he became World Champion in 2017 (individual and team TT) and won a silver medal at the 2016 Olympics. Further, the athlete won 2 GCs in the World Tour, a 7-day race the BinckBank Tour in 2017 and the Giro d’Italia in 2017. Until 2019, the athlete won 19 races including TTs and road races in all the GTs.

Research design
During 4 GTs, PO data were collected from as many as possible stages and uploaded to a central database. Because of sponsor changes, the collection of the PO data was done with different brands of power meters and with different power meters on different bikes (i.e. 1 road-bicycle, 2 reserve road bicycles and 2 TT bicycles). The PO data from the Vuelta a España 2015 and the Giro d’Italia 2017 was collected by Pioneer power-meters (SGY-PM910H2, Pioneer, Kawasaki, Japan) and from the Giro d’Italia 2018 and the Tour de France 2018 by Shimano power meter (FC-RC9100-P, Shimano, Sakia, Japan). The athlete and mechanics were informed about the importance of the zero calibration and were instructed to do the zero calibration before every ride. Due to malfunctions, crashes and bicycle changes PO data was not recorded for 2 stages at the Vuelta a España 2015 (i.e. stages 9 and 19), 2 stages at the Giro d’Italia 2017 (i.e. stages 3 and 4), 2 stages at the Giro d’Italia 2018 (i.e. stages 4 and 21) and 2 stages at the Tour de France 2018 (i.e. stages 2 and 13). All collected PO data were manually checked, and spikes were manually corrected when necessary. Data were sampled at 1 Hz during the Vuelta a España 2015, the Giro d’Italia 2017 and the Tour de France 2018. Using another bicycle computer during the Giro d’Italia 2018 the data were collected at 0.2
Hz. All bicycle computers measured race characteristics (i.e. distance, duration and elevation gain) and environmental characteristics (i.e. temperature).

**Load and intensity characteristics**

Load characteristics are based on PO and measured as the total mechanical energy spent (in kJ spent) and Training Stress Score (TSS). TSS was calculated according to Equation 2.1.

(Equation 2.1)

\[
TSS = \left( \frac{t \times NP \times IF}{FTP \times 3600} \right) \times 100
\]

Where \(t\) is the duration of the race in seconds and \(IF\) the intensity factor (see Equation 2.3). \(NP\) is the normalized power as calculated with Equation 2.2, where \(Pi\) is the floating mean power for 30 seconds time segments and \(N\) is the total number of time segments. The functional threshold power (FTP) was determined as 95% of the highest 20 minutes mean maximum PO obtained in races from that particular season. FTP was established at 408 W (5.8 W·kg\(^{-1}\)) for the Vuelta a España 2015, 409 W (5.8 W·kg\(^{-1}\)) for the Giro d’Italia 2017, 417 W (5.9 W·kg\(^{-1}\)) for the Giro d’Italia 2018 and 417 W (6.0 W·kg\(^{-1}\)) for the Tour de France 2018.

(Equation 2.2)

\[
NP = \sqrt[4]{\frac{1}{N} \sum_{i=1}^{N} P_i^4}
\]

(Equation 2.3)

\[
IF = \left( \frac{NP}{FTP} \right)
\]

In addition, similar to previous research, load metrics (TSS and kJ spent) are expressed relatively per kilometer (kJ spent·km\(^{-1}\) and TSS·km\(^{-1}\)). Intensity distribution was quantified based on the time spent in 5 different PO zones. The
5 PO zones were based on a percentage of FTP based on guidelines provided by Coggan\textsuperscript{23}: zone 1: $\leq55\%$ of FTP, zone 2: 56-75\% FTP, zone 3: 76-90\% FTP, zone 4: 91-105\% FTP, zone 5: $\geq106\%$ FTP.

**Performance characteristics**

To assess the performance during the 4 GTs, the MPP and the relative PO on the key mountains were analyzed. The MPP corresponded to the highest mean maximal power developed in each GT by the athlete for the durations of 5, 10, 30 seconds and 1, 5, 10, 20, 60, 120, 180 minutes. The MPP was expressed in relation to body mass of the cyclist ($W\cdot kg^{-1}$) to compensate for changes in body mass. Body mass was measured by the team for the four GTs and the mean body mass was 69.9 kg, 70.3 kg, 70.5 kg and 69.0 kg for the Vuelta a España 2015, the Giro d’Italia 2017, the Giro d’Italia 2018 and the Tour de France 2018, respectively.

To analyze the performance on the key mountains in the GTs, the last mountain in a stage with an uphill finish and mountains with a significant importance (e.g., finish directly after descend) were selected by the use of www.touretappe.nl\textsuperscript{77} and www.procyclingstats.com.\textsuperscript{78} Based on visual inspection of the PO, speed and altitude data, the key mountains were manually selected and PO data were collected and expressed in relation to the athlete’s body mass ($W\cdot kg^{-1}$).

**Factors that influence performance**

To investigate the effect of long-term fatigue on the climbing performance at the key mountains, we reported on which race-day the key mountains were in the GT (stage number). To determine the effect of short-term fatigue during the stage, the duration, distance, Total Elevation Gain (TEG), kJ spent, TSS, kJ spent·km$^{-1}$ and TSS·km$^{-1}$ were measured from the start of the stage till the start of the key mountain. Furthermore, to investigate the influence of mountain characteristics on relative PO we measured the duration and the gradient of the key mountains. Finally, temperature measured by the bike computer on the key mountain was analyzed to investigate the effect of heat stress on climbing performance.
Statistical analysis

Descriptive data are reported as mean (±SD). All parameters except stage number and temperature were collected by analyzing the collected PO data from the 4 GTs. Differences between the 4 GTs were determined by using one-way analysis of variance (ANOVA). Bonferroni’s post-hoc test was applied to identify differences when the ANOVA indicated a significant main effect. Stepwise Multiple Linear Regression analysis (SMLR) was used to identify the best predictors of the relative PO on the key mountains. Prior to SMLR, the collinearity between variables was determined by Pearson Correlations. For variables with a correlation of \( r > 0.7 \), the variable with the highest correlation with performance (relative PO on the key mountain) was used for SMLR. Based on a visual inspection of the relation between duration of the climb and relative PO, duration was logarithmically transformed for better fitting the data. Regression coefficients (intercept and slopes) are presented and uncertainties in the coefficients are presented as 95% Confidence Intervals (CI). PO analysis was performed using Golden Cheetah (Golden Cheetah, Version 3.4) and statistical analysis were performed using SPSS (IBM SPSS Statistics version 23, IBM Corporation, Armonk, NY, USA). The level of statistical significance was set at \( p < 0.05 \). The following criteria were adopted to interpret the magnitude of the correlation (\( r \)) between the measures: < 0.1 trivial, 0.1-0.3 small, 0.3-0.5 moderate, 0.5-0.7 large, 0.7-0.9 very large, and 0.9-1.0 almost perfect.\(^7\)

2.3 Results

In total, 76 stages were analyzed, which were collected during the Vuelta a España 2015 (n=19), the Giro d’Italia 2017 (n=19), the Giro d’Italia 2018 (n=19) and the Tour de France 2018 (n=19). In the GCs the athlete finished 6\(^{th}\), 1\(^{st}\), 2\(^{nd}\) and 2\(^{nd}\), respectively. In total the athlete finished in 14 stages in the top 3, from which he won 6 stages (2 road races and 4 TTs) and was GC-leader for 5 days, 9 days and 1 day in the Vuelta a España 2015, the Giro d’Italia 2017 and the Giro d’Italia 2018, respectively.

Table 2.1 presents basic descriptive and load characteristics for the four GTs. No significant differences were observed between the four GTs. The GT with the lowest load recorded was the Tour de France 2018 and the highest
load recorded was the Giro d’Italia 2018, the difference between both was 6% and 11% for kJ spent and TSS, respectively.

**Table 2.1:** Average load and intensity characteristics for the grand tours where the athlete competed for the general classification victory. Values are mean±SD.

<table>
<thead>
<tr>
<th></th>
<th>Vuelta a España 2015 (n stages = 19)</th>
<th>Giro d’Italia 2017 (n stages =19)</th>
<th>Giro d’Italia 2018 (n stages =19)</th>
<th>Tour de France 2018 (n stages =19)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Distance (Km)</strong></td>
<td>168 (53.5)</td>
<td>181 (55.1)</td>
<td>176 (58.2)</td>
<td>167 (55.4)</td>
</tr>
<tr>
<td><strong>Duration (Min)</strong></td>
<td>260 (85)</td>
<td>272 (95)</td>
<td>266 (91)</td>
<td>255 (86)</td>
</tr>
<tr>
<td><strong>PO (W)</strong></td>
<td>227 (36.2)</td>
<td>203 (43.2)</td>
<td>238 (47.2)</td>
<td>228 (36.5)</td>
</tr>
<tr>
<td><strong>PO (W·kg⁻¹)</strong></td>
<td>3.25 (0.52)</td>
<td>2.89 (0.61)</td>
<td>3.38 (0.67)</td>
<td>3.30 (0.53)</td>
</tr>
<tr>
<td><strong>TEG (m)</strong></td>
<td>2413 (1397)</td>
<td>2583 (1438)</td>
<td>2245 (1478)</td>
<td>2076 (1436)</td>
</tr>
<tr>
<td><strong>Total kJ spent (kJ)</strong></td>
<td>3623 (1196)</td>
<td>3643 (1185)</td>
<td>3770 (1359)</td>
<td>3530 (1192)</td>
</tr>
<tr>
<td><strong>Total TSS™</strong></td>
<td>233 (80.5)</td>
<td>237 (84.9)</td>
<td>234 (91.6)</td>
<td>209 (74.6)</td>
</tr>
<tr>
<td><strong>kJ spent · km⁻¹</strong></td>
<td>22.4 (4.8)</td>
<td>21.4 (5.9)</td>
<td>24.0 (11.7)</td>
<td>22.7 (7.6)</td>
</tr>
<tr>
<td><strong>TSS · km⁻¹</strong></td>
<td>1.48 (0.46)</td>
<td>1.48 (0.75)</td>
<td>1.39 (0.38)</td>
<td>1.38 (0.60)</td>
</tr>
</tbody>
</table>

Abbreviations: PO, power output; Norm, normalized; TEG, Total Elevation Gain; TSS, Training Stress Score. No Significant difference between Grand Tours are found (P > 0.05).

Figure 2.1 presents the relative and absolute time spent in the 5 PO zones. Post hoc test analyses revealed a lower absolute time spent in PO zone 5 during the Giro d’Italia 2018 (334±8 min) compared to the Vuelta a España 2015 (531±9 min) (P=0.005), the Giro d’Italia 2017 (531±11 min) (p=0.005) and the Tour de France 2018 (493±9 min) (P=0.039). No significant differences were observed between the other PO zones in the four GTs.

Figures 2.2, 2.3 and Table 2.2 give a detailed overview of the athlete’s performance during the four GTs. Figure 2.2 presents the relative MPP (5, 10, 30 seconds and 1, 5, 10, 20, 60, 120, 180 minutes) and figure 2.3 presents the relative PO of all key mountains. Table 2 provides a summary of the key mountains and the mountain characteristics. The average relative PO on all
The key mountains for all four GTs combined was 5.9±0.6 W∙kg⁻¹.

**Figure 2.1:** Intensity distribution expressed as relative (A) and absolute (B) time spent in different power output zones in a general classification contender during different Grand Tours. Absolute time spent in PO zone 5 during the Giro d’Italia 2018 is significantly (P < 0.05) different from to the other Grand Tours. PO indicates power output (w).

**Figure 2.2:** Maximum power profile from a general classification contender during different Grand Tours. The values indicate the maximum power output for different durations (5, 10, 30 seconds and 1, 5, 10, 20, 60, 120, 180 minutes) achieved in the grand tours participated by the athlete and are expressed in relation to the body mass of the athlete.

**Figure 2.3:** Relative power output at different key mountains in a general classification contender during multiple Grand Tours (Alto de la Mesa; Alto de Cazorla; Alto de Capileira; Alto Els Cortals d’Encamp; Alto Campo; Alto de Sotres; Alto Ermita de Alba; Puerto de Cotos; Blockhaus; Oropa; Selvino; Umbrailpass; Pontives; Piancavallo; Asiogo; Etna; Montevergine di Mercogliano; Gran Sasso d’Italia; Osimo; Zoncolan; Sappada; Pratonevoso; Bardonecchia; Cervinia; Mur d Bretagne; Col de la Colombiere; La Rosiere; Alpe d’Huez; Mende; Pic du Nore; Col du Portillon; Saint-Lary Soulan; Col d’Aubisque).
Table 2.3 presents the correlation matrix between all the factors with a moderate correlation with relative PO on the key mountain or factors identified by SMLR. SMLR analysis showed that in the final regression model climbing performance was significantly associated with a combination of the duration (log10 transformed) \((-1.52 \, \text{W} \cdot \text{kg}^{-1} \ [95\% \, \text{CI} -1.81 \text{ to} \ -1.22], \ r^2 = 0.61, \ p=0.0001)\), the TEG before the mountain \((-0.23 \, \text{W} \cdot \text{kg}^{-1} \ [95\% \, \text{CI} -0.32 \text{ to} \ -0.14], \ r^2 = 0.13, \ p=0.0001)\) and gradient of the mountain \((0.12 \, \text{W} \cdot \text{kg}^{-1} \ [95\% \, \text{CI} 0.07 \text{ to} \ 0.17], \ r^2 = 0.11, \ p=0.0001)\), while the intercept was \(7.64 \, \text{W} \cdot \text{kg}^{-1} \ [95\% \, \text{CI} 7.11 \text{ to} \ 8.17]\). A total 86% of the variance of the relative PO on the key mountain can be explained by those three factors, described in Equation 2.4.

\[
PO \ (\text{W} \cdot \text{kg}^{-1}) = 7.64 + \log_{10}(Duration \text{ mountain} \ (\text{min})) \times -1.52 + \text{Gradient Mountain} \% \times 0.12 - (TEG \text{ before Mountain} \ (\text{m}) \times 10^{-3}) \times 0.23
\]
**Table 2.3:** Correlation matrix for parameters with a moderate (r>0.3) association with performance on the key mountain or with a significant influence on performance based on stepwise multiple regression analysis.

<table>
<thead>
<tr>
<th>Performance</th>
<th>Mountain Characteristics</th>
<th>Long term fatigue</th>
<th>Short term fatigue</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PO (W·kg⁻¹)</strong></td>
<td><strong>Duration (min)</strong></td>
<td><strong>Gradient (%)</strong></td>
<td><strong>Stage number</strong></td>
</tr>
<tr>
<td>Performance</td>
<td>PO (W·kg⁻¹)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mountain Characteristics</td>
<td>Duration (min)</td>
<td>-0.78</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Gradient (%)</td>
<td>0.27</td>
<td>0.11</td>
</tr>
<tr>
<td>Long term fatigue</td>
<td>Stage number</td>
<td>-0.44</td>
<td>0.17</td>
</tr>
<tr>
<td>Short term fatigue</td>
<td>TEG</td>
<td>-0.45</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>kJ spent · km⁻¹</td>
<td>-0.40</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>TSS · km⁻¹</td>
<td>-0.40</td>
<td>-0.01</td>
</tr>
</tbody>
</table>

Performance, the relative PO on the last mountain; Mountain characteristics, duration and gradient of the last mountain; Stage-number, amount of days in the Grand Tour; Short-term fatigue, TEG, kJ spent · km⁻¹ and TSS · km⁻¹ measured before the last mountain. Abbreviations: PO, Power Output; TEG, Total Elevation Gain; kJ, kilojoule; TSS, Training Stress Score. * indicate a moderate or higher collinearity.
2.4 Discussion

This is, to the author’s knowledge, the first study describing the individual load, intensity and performance characteristics of a GC contender during multiple GTs. The presented data covers one GT in which the athlete finished 1<sup>st</sup> and two GTs where the athlete finished 2<sup>nd</sup>. Therefore, the present study gives an unique insight into the load, intensity and performance characteristics of the fight for a GT victory. Furthermore, SMLR analysis showed that the climbing duration and the gradient of the key mountain combined with indicators of short-term fatigue determined 86% of the variance of the relative PO on the key mountain.

Load and intensity

An impressive energy expenditure during racing between 74123 kJ and 79166 kJ or a TSS between 4400 au to 4983 au is necessary to finish a GT in the top-ten. Similar to Lucia et al.<sup>22</sup> this study did not find any significant differences between the load of the four GTs despite the fact that all GTs have different combinations of flat, semi-mountain, mountain stages and (team) TTs. The largest load differences occurred between the Giro d’Italia 2018 and Tour de France 2018, which was 6% and 11% for kJ spent and TSS, respectively. The differences between kJ spent and TSS are probably caused by the quadratic relation with exercise intensity which is integrated within the calculation of TSS.<sup>80</sup> Our data showed that despite competing for the GC, ~80% of the time in a GT is spent in the low-intensity zones (PO zones 1, 2 and 3). Further ~10% is spent around FTP (PO zone 4) and ~10% is spent in the highest PO zone or above FTP (PO zone 5). The only significant difference found between the four analyzed GTs was a lower absolute time spent in PO zone 5 during the Giro d’Italia 2018 (334±8.3 min) compared to the other three GTs. The lack of significant different load and intensity between the four analyzed GTs could suggest that GC contenders unconsciously pace their efforts during a GT.<sup>22,75</sup> For example, the Giro d’Italia 2017 had the highest amount of distance, duration and TEG while the PO, intensity factor and kJ spent·km<sup>-1</sup> were the lowest of all the four studied GTs. The lower time spent in PO zone 5 during the Giro d’Italia 2018 could be caused by the higher overall pace during this GT. This is shown by the higher PO and intensity factor. This is strengthened
by the results of the regression analysis that TEG before the key mountain had a decremental effect on relative PO. A high amount of elevation gains in a GT results in a lower PO on the key mountains and thus a lower load and intensity on the key mountains. This will result in an overall lower load and intensity demand for the specific GT. This is in agreement with Lucia et al., who stated that a shorter or less mountainous GT is compensated by a higher relative intensity compared to a GT which covers a lot of kilometers or altitude meters.

**Performances**

Reported MPP values in this study are higher compared to the values reported by Sanders et al., which is probably caused by averaging the values of a whole team (Figure 2.2). Cyclists have different roles during a GT (e.g. domestique, sprinter, GC contender and TT specialist) and different roles mean different priorities. For example, a GC contender must perform at his maximum on some moments during all the 21 stages, while sprinters and their domestiques pace mountain stages to be within the time limit and therefore don’t have to compete maximal every stage. Thus, averaging MPP values of GC contender and domestiques will result in lower MPP values compared to MPP values of a GC contender alone. In agreement with Pinot & Grappe the term “record power output” is preferred to describe performance instead of “mean maximum power output”, because the highest PO obtained during competition is not the maximum that can be achieved by the athlete. This could also be the reason why the values reported in this study are slightly lower compared to Pinot et al. The MPP values presented in this study were affected by fatigue, tactical racing and the ability to achieve a maximal effort exactly for each duration, while MPP values achieved during training are not influenced by the factors mentioned above (Figure 2.2). For the first time, climbing performances of a GC contender in a GT are described. On average an impressive 5.9±0.6 W·kg⁻¹ for 27±13 minutes is necessary on the key mountains to compete for the GC victory. The presented data is from a GC contender specialized in TTs and thus he gained time on his direct competitor (i.e. 1st or 2nd place in the GT) in the specific GTs during those TTs (i.e. 113 seconds, 257 seconds, 50 seconds and 14 seconds, in the respective GTs). Therefore, this GC contender can have a more defensive race strategy throughout the mountain stages, because his strategy is not to lose time in these stages. Thus, it could be that the described climbing
values are even higher in GC contenders specialized in climbing. Furthermore, this GC contender was GC leader for multiple days in the Vuelta a España 2015 (5 days) and the Giro d’Italia 2017 (9 days), defending the GC lead could result in a different race strategy and thus a different MPP or intensity distribution compared to a rider challenging the race leader.

Surprisingly, during the GT won by this athlete, the Giro d’Italia of 2017, the athlete had a slightly lower performance with 5.7±0.4 W·kg⁻¹ on the key mountains compared to the other GTs (Table 2.2). This could be explained by the results of the SMLR analyses. During this GT the average duration on the key mountains was longer. As shown by SMLR analysis, duration has a large negative influence on the relative PO on the key mountains. It was expected that the duration of the mountain has a large influence on climbing performance as it is well established that PO is influenced by duration described by the power duration relationship. This power duration relationship is also clearly visible in the MPP (Figure 2.2). Surprisingly, from all the parameters to assess short term fatigue, TEG before the key mountain has the largest influence on the performance on the key mountains, while load measurements before the key mountain (i.e. TSS and kJ spent) did not have any significant influence. However, load values expressed per kilometer (i.e. TSS·km⁻¹ and kJ spent·km⁻¹) before the key mountain did show a moderate relationship with performance on the key mountains (Table 2.3). In addition, collinearity analysis showed that they could be interchangeable used for TEG before the key mountain in the SMLR analysis. Thus TSS·km⁻¹ and kJ spent·km⁻¹ may be good indicators for short-term fatigue. One of the reasons that relative load values (TSS·km⁻¹ and kJ spent·km⁻¹) are better indicators for short-term fatigue compared to absolute load values. Professional cyclists are highly endurance trained and a long low-intensity race will not cause the same amount of fatigue as a shorter high-intensity race. The load measurements relative to distance (TSS·km⁻¹ and kJ spent·km⁻¹) are in line with previously reported measurements.

Further, gradient had a large effect on the relative PO on the key mountain, 14.5% of variance of the PO was determined by the gradient of the mountain. A 1% increase in steepness, means 0.12 W·kg⁻¹ higher relative PO on the mountain. This is in agreement with Padilla et al. who divided mountains during the three GTs into three categories based on the length and steepness and found higher PO on the steeper category mountains compared
to the less steeper category mountains. One of the reasons could be that a steeper mountain means less drafting and thus less help from domestiques. Therefore, gradient of the mountain also influences race-tactics which indirectly influence PO. Further a mountain with a lower gradient could mean some flatter easier parts and thus a more stochastic PO as drafting behind domestiques has a bigger effect with high speeds.

One of the hypotheses of this study was that long-term fatigue (stage number) would influence the performances on the key mountain in a GT. However, stage number did not significantly \((p=0.12)\) influence relative PO on the key mountain. This is not in line with previous research which found a decrement of performances after a GT of \(~10\%\) comparing a laboratory exercise test before and after a GT in non-GC contenders.\(^6\) Although not significant, the SMLR analyses indicate a PO decrement of \(-0.015 \text{ W} \cdot \text{kg}^{-1}\) per day in a GT which means a difference of \(~5\%) between a performance on the first day and the last day of a GT. One of the specific determinants of a successful GC contender is the ability to sustain accumulating load while holding a high-performance level. In the study of Rodriguez-Morrayo et al.\(^6\) cyclists with different specialties (none GC contenders) were studied, which could be a reason for the lower decrement in performance described in this study. Furthermore, a GC contender is protected by domestiques throughout the whole GT to save energy. In addition, we assume that with a higher number of climbs, long-term fatigue will be of significant influence on the performance on the key mountain. Another reason that stage number did not significantly influence performance is that most GTs have their mountain stages in week 2 or week 3 and therefore almost all performances analyzed in this study were already affected to some extent by long-term fatigue. In addition, it could be that fatigue influences the race intensity before the key mountain and thus had a smaller effect on the performance on the last mountain.

Our hypothesis was that a high environmental temperature would negatively influence performance on the key mountain because it is well known that heat stress impairs performance.\(^7\) This study did not find any significant influence of temperature on performance on the key mountain. The reason for this could be that only 4 mountains were recorded with a temperature above 30 degrees. From those 4 mountains, 2 mountains were shorter than 10 minutes and thus heat stress would probably not influence the
performance on those 2 mountains.

**Limitations**

The presented data were primarily collected for monitoring of training load of the athlete and are therefore not without limitations. The measurements of PO were done with 2 different brands of power meters and four different brands of bicycle computers. Also, the athlete had multiple bikes (1 road-bicycle, 2 reserve road bicycles and 2 TT bicycles) during the GTs and, therefore, the presented data of one GT were collected with different power meters. The mechanics of the team had the task to zero calibrate the power meter every day before the race although the authors did not control this. Further within the team no in house calibration was performed after receiving the power-meter from the manufacturers. Due to malfunctions, crashes and bicycle changes, PO was not measured during 8 stages divided equally over the 4 GTs. In addition, FTP was obtained by taking 95% of the highest 20 minutes mean maximum PO obtained during the particular season. Nimmerichter et al.\textsuperscript{83} showed that variations within a season can be up to 0.4 W·kg\textsuperscript{-1}. Furthermore the correction factor of 95% is somewhat arbitrary choices and it is shown that a low or high anaerobic capacity has an influence on this correction factor.\textsuperscript{84} Both limitations of FTP could had an influence on the presented TSS values and intensity zones. Furthermore, ideally, PO intensity zones would be anchored around physiological thresholds, such as the first and second lactate or ventilatory thresholds.\textsuperscript{24,85} However, during the time of the data collection no regular and controlled laboratory exercise testing was implemented within the team. Therefore, in this study the boundaries from the intensity zones are based on Coggan and Allen,\textsuperscript{23} which are somewhat arbitrary chosen and the boundaries are not equally spaced, this could influence the reported intensity distribution. Lastly, the key mountains were manually selected based on visual inspection of PO, speed and altitude profile and could therefore be slightly different in length compared to the official length of the mountains.
Conclusion
To conclude, overall load and intensity characteristics in four different GTs did not differ when competing for the GC. An impressive 5.7 to 6.0 W·kg⁻¹ on the key mountains is necessary to compete for the victory in multiple GTs. Short-term fatigue in combination with the gradient and duration of the climb determine 86% of the variance of the relative PO on the key mountains.

Acknowledgments
No funding is used for this research. We would like to thank the cyclist for his participation in this investigation. The results of the study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation. The results of the present study do not constitute endorsement by the American College of Sports Medicine.
Chapter 3

Intensity and Load Characteristics of Professional Road Cycling: Differences Between Men’s and Women’s Races.

Dajo Sanders • Teun van Erp • Jos J. de Koning

Abstract

Purpose: This study aims to provide an in-depth analysis of a large competition database describing the intensity and load demands of professional road cycling races, highlighting the differences between men’s and women’s races.

Method: In total, 20 male and 10 female professional cyclists participated in this study. During 4 consecutive years, heart rate, rating of perceived exertion and power output data were collected during both male (n = 3024) and female (n = 667) professional races. Intensity distribution in 5 HR zones was quantified. Competition load was calculated using different metrics including Training Stress Score (TSS), Training Impulse (TRIMP) and session rating of perceived exertion (sRPE). Standardized effect size is reported as Cohen’s d.

Results: Large to very large higher values (d = 1.36 – 2.86) were observed for distance, duration, total work (kJ) and mean power output in men’s races. Time spent in high intensity heart rate zones (i.e. zone 4 and zone 5) was large higher in women’s races (d = 1.38 – 1.55) compared to men’s races. Small higher absolute loads were observed in men’s races quantified using TSS (d = 0.53) and TRIMP (d = 0.23). However, load metrics expressed per kilometer was large to very largely higher in women’s races for TSS·km⁻¹ (d = 1.50) and TRIMP·km⁻¹ (d = 2.31).

Conclusions: Volume and load are higher in men’s races, whereas intensity and time spent at high-intensity is higher in women’s races. Coaches and practitioners should consider these differences in demands in the preparation of professional road cyclists.
3.1 Introduction

There are road cycling competitions all around the world across a broad spectrum that ranges from youth and junior competitions to elite professional competitions. A male World Tour professional cyclist will cycle around 30,000 to 35,000 km in training and competition each year including up to 100 competition days.\(^2,\(^{86}\) In recent years, women’s road cycling has been growing widely and in 2016 the Women’s World Tour was established with more and more races running alongside the Men’s World Tour program. Although published data on the training characteristics of female professional cyclists is limited, own observations have shown that female World Tour female cyclists will typically cover around 13,000 to 18,000 km in training and competition each year, including up to 65 competition days.

Due to technological advancements in recent years with mobile Heart Rate (HR) and power meters, the collection of both physiological (i.e. HR) and work rate (i.e. Power Output [PO]) data in the field is now widely possible to monitor the training and competition of cyclists. As a result of this accessible data collection, both applied and more descriptive studies on professional cycling (races) have been performed in recent decades. Most studies have focused on describing the demands of men’s professional road cycling races\(^37,\(^6,\(^{86-88}\) with limited research available describing the demands of women’s professional races.\(^42,\(^{89}\) However, even though some evidence regarding the demands of professional men’s and women’s races is available, there is little research describing the differences between men’s and women’s races in terms of exercise intensity and load demands. A detailed description of the demands of both men’s and women’s races is valuable information for coaches and practitioners working with these athletes on a daily basis. In addition, differences in intensity distribution between men’s and women’s races may result in different training prescription and preparation for races.

Accordingly, this study aims to provide an analysis of a large competition database describing the intensity and load demands of professional road cycling races, highlighting the differences between men’s and women’s races.
3.2 Methods

Participants
In total, 20 male (mean [SD]: age: 27.5 [4.0] y, height: 184.8 [6.2] cm, bodyweight: 73.2 [7.1 kg] and ten female (age: 24.5 [4.5] y, height: 169.6 [6.7] cm, bodyweight: 60.5 [4.3 kg] highly trained professional cyclists, part of a current World Tour professional cycling team, participated in this study. During the 4-year monitoring period, the men’s team were active on the Pro-Continental level for the first year and part of the World Tour for the last three years. The women’s team finished every year within the top-10 of the UCI elite women team ranking over the course of the study period. Institutional ethics approval was granted and, in agreement with the Helsinki Declaration, written informed consent was obtained from the participants.

Research design
During 4 consecutive years, Rating of Perceived Exertion (RPE), HR and PO data was collected during both single-day and multi-day (stage) races for the males and females on the team. Depending on how long the cyclist was involved in riding for the team, the data set of an individual cyclist contains data ranging from 1 to 4 years. If a cyclist was not able to ride for a period of 3 months or more, because of illness or an injury, the data set of this particular year was excluded. All data sets were visually checked for erroneous data and incomplete data files due to technological issues (e.g. flat battery of power meter or monitor) were excluded. If one of the three main variables (i.e. RPE, HR or PO) was missing for a given race, and no erroneous data was present within the given file, this dataset was still analyzed using the available data.

Race characteristics
Intensity distribution was quantified based on the time spent in different HR zones. A 5-zone model was used to quantify the time spent in each intensity zone. HR zones were based on percentages of maximal heart rate (HR$_{\text{max}}$) (zone 1: 50-59% HR$_{\text{max}}$, zone 2: 60-69% HR$_{\text{max}}$, zone 3: 70-79% HR$_{\text{max}}$, zone 4: 80-89% HR$_{\text{max}}$, zone 5: 90-100%). HR$_{\text{max}}$ was defined as the highest HR achieved by the cyclist during training or competition of the analyzed season and adjusted every season (if needed). The determination of intensity
zones is ideally approached using the integration of physiological measures and anchored around physiological thresholds (i.e. lactate or ventilatory thresholds), however, as no structured laboratory exercise testing was incorporated over the course of this study, this was not feasible in this study. In addition, the percentage of total race time spent at different power bands was compared between men’s and women’s races. The power bands were constructed in steps of 0.75 W·kg⁻¹ ranges from < 0.75 to > 7.50 W·kg⁻¹.

Exercise load was calculated using different methods based on either HR, PO or RPE: Edwards’ training impulse (eTRIMP), Training Stress Score (TSS) and session RPE (sRPE). Edwards’ TRIMP was calculated based on the time spent in the 5 predefined HR zones described above and multiplied by a zone-specific arbitrary weighting factor (zone 1: weighting factor = 1, zone 2: weighting factor = 2, zone 3: weighting factor = 3, zone 4: weighting factor = 4, zone 5: weighting factor = 5) and then summated to provide a total eTRIMP score. TSS was calculated based on power data collected with portable power meters (SRM, Jülich, Welldorf, Germany and Pioneer, Kawasaki, Japan). TSS was calculated according to Coggan, using the following formula:

\[ TSS = \left( \frac{t \times NP \times IF}{FTP \times 3600} \right) \times 100 \]

Where \( t \) is the duration of the exercise bout in seconds, NP is Normalized Power of the exercise bout, and IF is an Intensity Factor that is the ratio between the NP of the exercise bout and the individual’s Functional Threshold Power (FTP). FTP was determined as 95% of the highest 20 min mean maximal PO, either achieved during a specific 20-minutes time trial in training or adjusted when the mean maximal 20 minutes PO was higher during a race. All riders were informed about the importance of the zero calibration of the power meter and were instructed to do the zero calibration before every ride. Both Edwards’ TRIMP and TSS have previously been shown to have a strong dose-response relationship with changes in fitness in competitive road cyclists. As a subjective measure of internal load, sRPE was calculated using the participants’ RPE (6-20 scale) and session duration. Riders were familiarized with the RPE scale prior to the start of this study and were instructed on the use of the scale. The RPE was obtained after the race, using
an online, self-filled in logbook, based on the question: “How hard was your workout?”. Even though the general recommendation is to obtain a RPE score within 30 minutes of each competition, the time between the end of the race and the cyclist filling in the RPE score could have been longer in this study (~1 – 5 hours). However, previous studies have shown that athletes are able to recall RPE accurately between 24 – 48 hours after the end of the training or competition. Exercise load for the session was then quantified by multiplying the RPE by the duration of the session (in minutes). In addition, similar to previous research, all load metrics (eTRIMP, TSS and sRPE) as well as total work performed (in kilojoules) were also expressed relatively per kilometer (i.e. eTRIMP∙km⁻¹, TSS∙km⁻¹, sRPE∙km⁻¹ and kJ∙km⁻¹).

**Statistical analysis**

Prior to analysis, the assumption of normality was verified by using Shapiro-Wilk test and by visual inspection of Q-Q plots. Intensity and load variables were compared to each other using a multilevel random intercept model using Tukey’s method for pairwise comparisons in R (R: A Language and environment for statistical computing, Vienna, Austria). Random effect variability was modelled using a random intercept for each individual participant. Level of significance was established at P < 0.05. In addition, magnitude based inferences was used to further evaluate and describe the magnitude of the effects observed. Standardized effect size is reported as Cohen’s d, using the pooled standard deviation as the denominator. Qualitative interpretation of d was based on the guidelines provided by Hopkins et al.: 0 to 0.19, trivial; 0.20 to 0.59, small; 0.6 to 1.19, moderate; 1.20 to 1.99, large; ≥ 2.00, very large.

### 3.3 Results

A total of 616 women’s races and 3024 men’s races were collected and analyzed. In total, 3640 races with power data (women’s races; n = 616, men’s races; n = 3024), 2346 races with HR data (women’s races; n = 424, men’s races; n = 1730) and 1621 races with RPE (women’s races; n = 533, men’s races; n = 1088) data were analyzed. The main part of the dataset included multi-day stage races for both men and women (78% of men’s races, 60% of women’s
Within the dataset, there were a total of 57 wins (1.9% of total files) and 289 top-10 finishes (9.6% of total files) for the men’s team and 7 wins (1.1% of total files) and 121 top-10 finishes (19.6% of total files) for the women’s team.

Table 3.1 presents the descriptive values for both the men’s and women’s races. Large to very large higher values (d = 1.36 – 2.86) were observed for distance, duration, mean PO and in men’s races. However, Intensity Factor™, mean HR and mean HR as %HRmax were largely higher (d = 1.36 – 1.80) in women’s races. Figure 3.1 graphically displays the differences in the percentage time spent in different HR zones for men’s versus women’s professional cycling races. Time spent in high intensity HR zones (i.e. zone 4 and zone 5) was largely higher in women’s races (d = 1.38 – 1.55) compared to men’s races (42% [11%] and 21 [16%] versus 24% [12%] and 6% [6%]).

Table 3.1: Volume and intensity characteristics of women’s and men’s professional cycling races.

<table>
<thead>
<tr>
<th></th>
<th>Women’s races</th>
<th>Men’s races</th>
<th>Differences men vs women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Cohen’s d</td>
</tr>
<tr>
<td>Distance (km)</td>
<td>116 (17)</td>
<td>183 (32) *</td>
<td>2.70</td>
</tr>
<tr>
<td>Duration (min)</td>
<td>194 (30)</td>
<td>285 (56) *</td>
<td>2.12</td>
</tr>
<tr>
<td>Mean PO (W)</td>
<td>167 (21)</td>
<td>216 (34) *</td>
<td>1.81</td>
</tr>
<tr>
<td>Mean PO (W·kg⁻¹)</td>
<td>2.8 (0.4)</td>
<td>3.0 (0.5) *</td>
<td>0.44</td>
</tr>
<tr>
<td>Intensity Factor™</td>
<td>0.83 (0.07')</td>
<td>0.73 (0.08) *</td>
<td>1.36</td>
</tr>
<tr>
<td>Mean HR (beats·min⁻¹)</td>
<td>152 (13)</td>
<td>133 (12) *</td>
<td>1.60</td>
</tr>
<tr>
<td>Mean HR (%HRmax)</td>
<td>79 (10)</td>
<td>69 (6) *</td>
<td>1.80</td>
</tr>
<tr>
<td>HRmax (beats·min⁻¹)</td>
<td>185 (10)</td>
<td>180 (12)</td>
<td>0.40</td>
</tr>
<tr>
<td>Mean RPE (AU)            (6-20 scale)</td>
<td>15.4 (1.5)</td>
<td>15.4 (2.1)</td>
<td>0.02</td>
</tr>
</tbody>
</table>

**Abbreviations:** PO, power output; HR, heart rate; HRmax, maximal heart rate; RPE, rating of perceived exertion. Qualitative interpretation of d was based on the guidelines provided by Hopkins et al.95 *Significant difference (P < 0.05)

Table 3.2 presents the absolute and relative load of the men’s and women’s races. Absolute and relative total work was large to very largely (d = 1.48 – 2.73) higher in men’s races. Small higher absolute loads were observed in men’s races quantified using TSS (d = 0.53) and eTRIMP (d = 0.23). However, relative
load was *large to very largely* higher in women’s races for TSS∙km⁻¹ (d = 1.50) and eTRIMP∙km⁻¹ (d = 2.31).

**Figure 3.1:** Intensity distribution as %time spent in different heart rate zones in men’s versus women’s professional cycling races. HR indicates heart rate. Data are Mean(SD).

**Table 3.2:** Absolute and relative load in men’s versus women’s professional cycling races

<table>
<thead>
<tr>
<th></th>
<th>Women’s races</th>
<th>Men’s races</th>
<th>Differences men vs women</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total work (kJ)</strong></td>
<td>1958 (384) *</td>
<td>3734 (918)</td>
<td>2.73 – <em>Very large</em></td>
</tr>
<tr>
<td><strong>TSS (AU)</strong></td>
<td>224 (49) *</td>
<td>255 (68)</td>
<td>0.53 – <em>Small</em></td>
</tr>
<tr>
<td><strong>TRIMP (AU)</strong></td>
<td>700 (141) *</td>
<td>739 (203)</td>
<td>0.23 – <em>Small</em></td>
</tr>
<tr>
<td><strong>sRPE (AU)</strong></td>
<td>2982 (585) *</td>
<td>4370 (1144)</td>
<td>1.61 – <em>Large</em></td>
</tr>
<tr>
<td><strong>kJ spent ∙ km⁻¹</strong></td>
<td>16.8 (2.5) *</td>
<td>20.4 (3.9)</td>
<td>1.48 – <em>Large</em></td>
</tr>
<tr>
<td><strong>TSS ∙ km⁻¹</strong></td>
<td>1.92 (0.35) *</td>
<td>1.40 (0.31)</td>
<td>1.50 – <em>Large</em></td>
</tr>
<tr>
<td><strong>TRIMP ∙ km⁻¹</strong></td>
<td>6.02 (0.84) *</td>
<td>4.08 (0.94)</td>
<td>2.31 – <em>Very large</em></td>
</tr>
<tr>
<td><strong>sRPE ∙ km⁻¹</strong></td>
<td>25.6 (4.0) *</td>
<td>24.0 (4.3)</td>
<td>0.42 – <em>Small</em></td>
</tr>
</tbody>
</table>

Abbreviations: TSS, Training Stress Score; TRIMP, Edwards’ training impulse; sRPE, session rating of perceived exertion. Note: Qualitative interpretation of Cohen’s d was based on the guidelines provided by Hopkins et al.95 *Significant difference (P < 0.05).
Figure 3.2 presents the percentage of competition time spent at different relative PO (W·kg⁻¹) bands. Time spent at the lower end of the power bands (0.76 – 3.00 W·kg⁻¹) was *moderately* higher for women’s races (d = 0.65 – 1.16). Time spent at the higher end of the power bands (4.51 – 6.75 W·kg⁻¹) was *moderately* higher for men’s races (d = 0.60 – 0.72).

![Figure 3.2: Power output distribution as % time spent in different power bands. Note: Qualitative interpretation of d was based on the guidelines provided by Hopkins et al. *presents a moderate difference (d ≥ 0.60). Data are Mean(SD).](image)

Table 3.3 presents the differences in intensity and load metrics between men’s and women’s races for both single-day and multi-day stage races. Mean HR, mean HR as a percentage of HRₘₐₓ and eTRIMP·km⁻¹ were *moderately* higher in single-day compared to multi-day races for both men and women (d = 0.66 – 0.96). Mean PO (W·kg⁻¹), RPE, Intensity Factor and TSS·km⁻¹ were also higher in single-day races for both men and women with these differences being *small* (d = 0.26 – 0.59).

### 3.4 Discussion

This study aimed to provide an in-depth analysis of a large competition database (~3700 professional road cycling races) describing the intensity and load demands of professional road cycling races, highlighting the differences between men’s and women’s races. This study reports the substantial differences in intensity and load characteristics of men’s versus women’s races. Within expectations, men’s cycling races are higher in duration, distance, total work, absolute PO and load. However, women spent a substantially bigger proportion of time at higher intensity zones compared to men’s races. In addition, relative load (TSS·km⁻¹ and eTRIMP·km⁻¹) is *large* to very *largely*
higher in women’s races compared to men’s races. These descriptive results contribute to a better understanding of the demands of professional cycling races and the specific differences between men’s and women’s races.

Because of the differences in race format and regulations in men’s versus women’s races, the substantial higher duration, distance and total work (in kilojoules) are not surprising. Following the regulations of the international cycling federation, the UCI, 1-day (professional) races for women are limited to a maximum of 160 km on the highest level (“World Tour”) whereas the longest one-day races for men can be around 260 even up to 300 km. Obviously, these regulations largely contribute to the observed differences within this study, especially relating to the ‘volume’ based metrics. However, metrics expressed

Table 3.3: Differences in intensity and relative load between women’s and men’s races for single-day and multiday races.

<table>
<thead>
<tr>
<th></th>
<th>Single day</th>
<th></th>
<th></th>
<th>Multiday</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Women’s</td>
<td>Men’s</td>
<td>Differences men vs women</td>
<td>Women’s</td>
<td>Men’s</td>
</tr>
<tr>
<td></td>
<td>races</td>
<td>races</td>
<td></td>
<td>races</td>
<td>races</td>
</tr>
<tr>
<td>Mean PO (W·kg⁻¹)</td>
<td>2.84 (0.30) *</td>
<td>3.17 (0.41)</td>
<td>0.94 – Moderate</td>
<td>2.68 (0.34) *</td>
<td>2.99 (0.43)</td>
</tr>
<tr>
<td>Mean HR (beats·min⁻¹)</td>
<td>157 (10) *</td>
<td>140 (10)</td>
<td>1.57 – Large</td>
<td>149 (12) *</td>
<td>130 (10)</td>
</tr>
<tr>
<td>Mean RPE (AU)</td>
<td>15.2 (1.6) *</td>
<td>16.1 (1.9)</td>
<td>0.51 – Small</td>
<td>15.5 (1.5) *</td>
<td>15.6 (2.0)</td>
</tr>
<tr>
<td>Intensity Factor™</td>
<td>0.86 (0.05) *</td>
<td>0.76 (0.07)</td>
<td>1.40 – Large</td>
<td>0.82 (0.08) *</td>
<td>0.73 (0.07)</td>
</tr>
<tr>
<td>Mean HR (%HRmax)</td>
<td>81 (4) *</td>
<td>74 (5)</td>
<td>1.77 – Large</td>
<td>77 (4) *</td>
<td>69 (5)</td>
</tr>
<tr>
<td>TSS (AU)</td>
<td>236 (48) *</td>
<td>286 (80)</td>
<td>0.78 – Moderate</td>
<td>217 (52) *</td>
<td>254 (63)</td>
</tr>
<tr>
<td>TRIMP (AU)</td>
<td>737 (151) *</td>
<td>868 (207)</td>
<td>0.73 – Moderate</td>
<td>670 (124) *</td>
<td>702 (186)</td>
</tr>
<tr>
<td>TSS·km⁻¹ (AU)</td>
<td>2.01 (0.30) *</td>
<td>1.49 (0.27)</td>
<td>1.80 – Large</td>
<td>1.88 (0.37) *</td>
<td>1.41 (0.30)</td>
</tr>
<tr>
<td>eTRIMP·km⁻¹ (AU)</td>
<td>6.33 (0.72) *</td>
<td>4.63 (0.72)</td>
<td>2.34 – Very large</td>
<td>5.80 (0.85) *</td>
<td>4.04 (0.90)</td>
</tr>
</tbody>
</table>

Abbreviations: HR, heart rate; HRmax, maximal heart rate; PO, power output; RPE, ratings of perceived exertion; TSS, Training Stress Score; eTRIMP, Edwards’ training impulse; sRPE, session rating of perceived exertion. Note: Qualitative interpretation of Cohen’s d was based on the guidelines provided by Hopkins et al. *Significant difference (P < 0.05).

Chapter 3

54
relatively (i.e. % of total race time, load per kilometer) were substantially higher in women’s races. This is nicely illustrated in Figure 3.1 showing a substantially higher proportion of time spent at the highest HR zones (i.e. zone 4 and zone 5) in women’s races compared to men’s races. In addition, the mean HR relative to their maximal HR is 10% higher (79% vs 69%) in women’s races compared to men’s races. Hence, it seems that women compensate the shorter duration of their races with a higher intensity and different riding style. To the best of the authors’ knowledge, this is the first study to specifically highlight the substantial differences in intensity demands comparing men’s to women’s professional cycling races. Due to the substantial differences in the proportion of total competition time spent at high intensity, training strategies to prepare for these demands (e.g. high intensity interval training formats) may not be used interchangeably for male and female professional cyclists.

It is important to note that, despite of the substantial differences in objective intensity metrics based on HR and PO between men’s and women’s races, perceived intensity (i.e. RPE) was not different when comparing men’s and women’s races. Hence, differences in physical demands, objective intensity and load characteristics may still result in similar perceived intensity, suggesting that subjective and objective metrics have the ability to reflect different constructs within the training process. This is in line with previous research in cyclists describing the differences between subjective and objective measurements of intensity (and load) in evaluating training characteristics and how the combination of subjective and objective metrics can be used to detect states of excessive fatigue or adaptation.27,36,96

Mean PO was largely higher in men’s races compared to women’s races which is suggested to be largely determined, among other factors, by inherent physiological differences between men and women, specifically relating to maximal oxygen uptake and body composition (i.e. higher lean mass in males).97 This results in the typically higher aerobic capacity observed in male professional cyclists56 compared to female cyclists.49 When controlled for bodyweight, the differences between men’s and women’s races in terms of mean PO is decreased; from a large difference (d = 1.81) for absolute PO to a small difference (d = 0.44) in relative PO (W·kg⁻¹). The substantial lower bodyweight (60.5 [4.3] kg vs 73.2 [7.1] kg) for the female cyclists in this study most likely contributes to the smaller difference. When comparing to previous
literature, the observed mean PO in this study during women’s races (2.8 [0.4] W·kg⁻¹) is lower compared to the 3.0 – 3.4 W·kg⁻¹ previously reported mean POs for women’s world cup races.42,89 The main reason proposed for this difference is that the previously reported values only evaluated world cup races whilst the competition database in this study also incorporated non-world cup (i.e. lower level) races in the analysis. In terms of men’s races, Ebert et al.87 showed a mean PO of 2.7 W·kg⁻¹ for flat and 2.9 W·kg⁻¹ for hilly professional male races87, whilst we observed a mean PO of 3.0 [0.5] W·kg⁻¹ in this study. Furthermore, the mean PO observed in this study is similar to what has previously been observed in professional cyclists during a multistage cycling race (3.1 [0.2] W·kg⁻¹)37 and the mean PO during the competitive season of 4 professional cyclists (3.1 W·kg⁻¹).86 Even though there are some discrepancies between studies, based on the current evidence, female professional cycling races will vary on average around 2.8 W·kg⁻¹ with world cup races > 3.0 W·kg⁻¹. On average, male professional cycling races will vary around 3.0 – 3.1 W·kg⁻¹ whilst this may be higher or lower depending on the level of competition, the race profile (e.g. elevation gain69,70) and race tactics.

Intensity and load demands of professional men’s cycling races has been evaluated in a number of previous studies,69,70,86,88 however, studies evaluating the characteristics of women’s professional cycling races remains limited. Recently, Menaspa et al.42 evaluated the demands of world cup competitions in professional women road cycling races. Although the reporting of their results doesn’t allow an exact comparison, percentage of competition time spent at different power bands seems to be similar in this study compared to the results by Menaspa et al.,42 a large proportion of competition time is spent at PO < 0.75 W·kg⁻¹ due to non-pedaling activity. Besides that, similar to the results by Menaspa et al.,42 a big proportion of time is spent at 1.51 – 4.50 W·kg⁻¹ in women’s races. A slight shift to the right can be seen in terms of the proportion of competition spent at the different power bands for men’s races with a big proportion of time spent at 2.26 – 5.25 W·kg⁻¹. This right shift in the power bands is caused by the higher relative and absolute PO of men’s races compared to women’s races (Table 3.1). Although this provides valuable insight in to the (mean) demands of professional cycling races, it should be acknowledged that the level of competition (i.e. World Tour vs no World Tour races), level of athlete,42 race profile70 and race tactics42 can
have a large effect on the quantified demands of the race. However, as the main aim of this study was to examine the differences in intensity and load of men's versus women's professional cycling as a whole – and not the differences between the demands of different levels of races, it was chosen to adopt an approach where all the data was analyzed and compared in order to maximize the sample size. Furthermore, the level hierarchy of professional men's races is more extensive and complicated compared to women’s cycling making such direct comparisons difficult to interpret.

In line with previous research,\textsuperscript{98} small to moderate higher intensity and relative loads were observed for single-day races compared to multi-day stage races for both men's and women's races. However, irrespective of the race format (i.e. single or multi-day race) intensity and relative load was higher in women’s races for both single and multi-day races. For example, during single-day and multi-day races mean HR was at 74\% and 69\% of HR\textsubscript{max} for men’s races, whereas it was at 81\% and 77\% of HR\textsubscript{max} for women’s races, respectively.

**Limitations**

There are some limitations that need to be taken into account when interpreting the results of the study. The main causes for limitations occurring with this analysis come from the fact that this was a retrospective analysis of race data. For example, the HR zones used in this study are based on ranges of a percentage of maximal HR whilst it must be acknowledged that there can be day-to-day variations in maximal HR (e.g. due to fatigue\textsuperscript{36,60}) that can influence the data analysis. Ideally, HR zones would be anchored around physiological thresholds such as the first and second lactate or ventilatory thresholds.\textsuperscript{24,25,90} However, during the time of the analysis, no regular and controlled laboratory exercise testing was implemented within the team, making such approaches (retrospectively) not feasible. Furthermore, for this same reason, more individualized approaches to load quantification previously used in cycling\textsuperscript{92} such as individualized TRIMP\textsuperscript{99} or Lucia's TRIMP\textsuperscript{22} were not feasible in our study. However, it must be noted that both Edward's TRIMP and TSS showed strong dose-response validity with changes in aerobic fitness in competitive road cyclists.\textsuperscript{92} In addition, FTP which was determined using the year's best 20 min mean maximal PO achieved in
training or racing. Hence, during certain time periods FTP can be either under- or overestimated which would lead to variability and inaccuracies with regards to the determination of TSS. With acknowledging these limitations caused by the retrospective analysis, this approach has made it possible to collect and analyze a large competition database in elite athletes (~3700 races), which has currently not been published before. Thereby, despite of the mentioned limitations, this study highlights important differences in competition intensity and load demands between men’s and women’s races.

**Practical Application**
These descriptive results contribute to a better understanding of the demands of professional cycling races and the specific differences between men’s and women’s races. Within expectations, men’s cycling races are higher in duration, distance, total work (kJ), absolute PO and load. However, the intensity of women’s races is substantially higher compared to men’s races highlighted by the time spent in high-intensity zones and the *large to very large* higher relative load in women’s races. Coaches and practitioners should consider these differences in demands in the preparation of professional cyclists. These results may indicate that preparation strategies between men’s and women’s races cannot be used interchangeably. The substantial differences in the time spent at high intensity (HR zones) in women’s races may suggest that preparation strategies for these demands (e.g. high intensity interval training formats/protocols) may require a different approach compared to men’s races.

**Conclusions**
To conclude, although overall volume and load characteristics are higher in men’s races, time spent at high intensity and relative load is higher in women’s races, despite a similar RPE. Coaches and practitioners should consider these differences in demands in the preparation of professional road cyclists. These results may indicate that preparation strategies between men’s and women’s races cannot be used interchangeably.
Acknowledgments
No sources of funding were used to compose this article. The authors have no conflicts of interest that are related to the described content of this manuscript.
Chapter 4

Training Characteristics of Male and Female Professional Road Cyclists: A Four Year Retrospective Analysis.

Teun van Erp • Dajo Sanders • Jos J. de Koning

Int J Sports Physiol Perform. [Epub ahead of print]
Abstract

**Purpose:** This study aims to describe the training intensity and load characteristics of professional cyclists using a four-year retrospective analysis. Particularly, this study aims to describe the differences in training characteristics between men and women professional cyclists.

**Method:** For 4 consecutive years, training data was collected from 20 male and 10 female professional cyclists. From those training sessions, Heart Rate (HR), rating of perceived exertion and Power Output (PO) was analyzed. Training intensity distribution as time spent in different HR and PO zones was quantified. Training load was calculated using different metrics such as Training Stress Score (TSS), Training Impulse (TRIMP) and session-RPE (sRPE). Standardized effect size is reported as Cohen’s $d$.

**Results:** *Small to large* higher values were observed for distance, duration, kJ spent and (relative) mean PO in men’s training ($d = 0.44 – 1.98$). Furthermore, men spent more time in low-intensity zones (i.e. zones 1 and 2) compared to women. *Trivial* differences in training load (i.e. TSS and TRIMP) were observed between men’s and women’s training ($d = 0.07 – 0.12$). However, load values expressed per kilometer were *moderately* ($d = 0.67 – 0.76$) higher in women compared to men’s training.

**Conclusions:** Substantial differences in training characteristics exist between male and female professional cyclists. Particularly it seems that female professional cyclists compensate their lower training volume, with a higher training intensity, in comparison to male professional cyclists.
4.1 Introduction

Professional road cycling is one of the most demanding sports in the world, where the longest 1-day race in men’s cycling can be up to 300 km and the longest multiple day stage races up to 21 days. In recent years professional female cycling has grown significantly, getting more and more attention from the public. The longest 1-day race in professional women’s cycling is 160 km and the longest multiple stage race is the Giro d’Italia Internazionale Femminile which contains 10 race-days.

The introduction of mobile Heart Rate (HR) monitors and power-meters makes the quantification of load and intensity characteristics of cyclist’s training and racing widely accessible. Alongside this, applied and descriptive studies are now published describing the load and intensity characteristics of professional cycling. However the main focus of the research published is on the load and intensity characteristics of men’s professional road cycling races and multiple day events, with some exceptions focusing on women’s races.

Recently Sanders et al. presented a retrospective analysis of a large competition database and reported the load and intensity characteristic of 616 women’s and 3024 men’s professional road races. They reported several substantial differences between the load and intensity demands of male races and female races. As expected, the absolute values for distance, total work (kJ spent) and absolute PO are higher for men’s races. However, load values expressed relative to distance or duration are large to very large higher in women’s races. Furthermore, the distribution of HR high-intensity zones (i.e. zones 4 and zone 5) are significantly different between men’s and women’s races. Female cyclists spent a larger proportion of their time in the high-intensity zones (zone 4: 42% vs 24% and zone 5: 21% vs 6%).

Subsequently, it was argued that training characteristics in women’s cycling may need to differ from men’s cycling to be optimally prepared for the different demands during races and that training strategies may not be used interchangeably between women and men professional cyclists. However, limited research is available describing the training characteristics of male professional cycling and to the best of our knowledge there is no research providing a detailed quantification of the training demands of
female professional cyclists. Therefore, this study evaluated the difference between training characteristics (quantified using both objective and subjective metrics) of both male and female professional cyclists using a 4-year retrospective analysis.

4.2 Methods

Research design
During 4 consecutive years, field data was collected within a professional cycling team with the aim of monitoring the cyclists training and performance. During the 4 years of data collection the women’s team finished each year within the top 10 of the Union Cycliste International (UCI) elite team ranking. During the first year of data collection the men’s team was active on Pro-Continental level and the last 3 years at World Tour level. From as much as possible training days HR, Power Output (PO) and Rating of Perceived Exertion (RPE) data were collected. An individual data-set could vary in length from 1 to 4 years depending on their respective lengths of stay within the team. If a cyclist was not able to ride for more than 3 months in 1 season, his or her data from that particular year was excluded from analysis. All data was visually checked for erroneous data or incomplete data files due to technological issues (e.g. flat battery of power-meter) were removed when necessary. If one of the 3 main variables were missing for a given training session and no erroneous data were present within the given session, the data set was still analyzed using the available variables.

Participants
In total 20 male and 10 female elite professional cyclists participated in this study. All cyclists were racing in the same professional cycling team. Men’s mean (SD) age: 27.5 (4.0) y, height: 184.8 (6.2) cm, body weight: 73.2 (7.1) kg and 20 minutes mean maximum power: 389 (27) W and women’s were 24.5 (4.5) y, height: 169.6 (6.7) cm, body weight: 60.5 (4.3) kg and 20 minutes mean maximum power 256 (19) W. Institutional ethics approval was granted and, in agreement with the Helsinki Declaration, written informed consent was obtained from the participants.
Training characteristics

To investigate the difference in the intensity distribution between the training of elite male and female cyclists, time spent in different HR and PO zone were quantified. A 5-zone model was used to quantify the time spent in each intensity zone. HR zones were based on percentage of maximal heart rate ($HR_{max}$) (zone 1: 50-59% $HR_{max}$, zone 2: 60-69% $HR_{max}$, zone 3: 70-79% $HR_{max}$, zone 4: 80-89% $HR_{max}$, zone 5: 90-100%). Maximal HR was defined as the highest HR ($HR_{max}$) recorded during every season (from November until October). The five PO zones were based on a percentage of Functional Threshold Power (FTP) based on guidelines provided by Coggan: $^{23}$ zone 1: ≤55% of FTP, zone 2: 56-75% FTP, zone 3: 76-90% FTP, zone 4: 91-105% FTP, zone 5: ≥106% FTP. FTP was determined as the highest 20 minutes mean maximal PO obtained for that specific season and adjusted accordingly every season. Further, according to previous research, $^{28,42}$ the percentage of time spent at different power outputs was calculated over 11 power bands with steps of 0.75 W·kg$^{-1}$, ranging from <0.75 W·kg$^{-1}$ to > 7.50 W·kg$^{-1}$.

To investigate the differences in male and female training load, 4 different measurements of training load were calculated. Internal training load was quantified using TRIMP$^{20}$ and sRPE$^{10,17}$ and external training load was quantified using kJ spent (mechanical energy spent) and TSS.$^{23}$ TRIMP was based on 5 HR-zones described above and multiplied by a zone-specific arbitrary weighting factor (zone 1: weighting factor = 1, zone 2: weighting factor = 2, zone 3: weighting factor = 3, zone 4: weighting factor = 4, zone 5: weighting factor = 5) and then summated to provide a total TRIMP score. TSS and kJ spent were calculated based on PO data collected by SRM and Pioneer power meters (SRM, Jülich, Welddorf, Germany and Pioneer, Kawasaki Japan). Riders were encouraged to perform the zero calibration before every ride. TSS was calculated according to Equation 4.1.$^{23}$

$$TSS = \left( \frac{t \times NP \times IF}{FTP \times 3600} \right) \times 100$$

Where $t$ is the duration of the session in seconds, where NP is the normalized power of the exercise,$^{23}$ and IF (intensity factor) is the ratio between NP and FTP.
kJ spent was calculated from the PO of every ride. RPE-scores were obtained by a daily online logbook, which all riders completed as soon as possible after training. One of the questions in the logbook was: How hard was your training? In which the cyclists would provide a RPE score to this questions based on a 6 to 20 BORG scale.\textsuperscript{18} The sRPE was calculated by multiplying the RPE of the session by the duration of the session in minutes, according to Foster et al.\textsuperscript{10,17} Although riders were instructed to fill in the RPE-score directly after training, there were instances where there was a slight delay in providing the score (e.g. the evening after the training session). However, previous studies have shown that RPE-scores are remarkably robust and athletes are able to recall RPE accurately after 24 hours after the end of their training or competition.\textsuperscript{100} In addition all parameters of training load (TRIMP, TSS, kJ spent and sRPE) were also expressed relatively per kilometer (i.e. TRIMP\(\cdot\)km\(^{-1}\), TSS\(\cdot\)km\(^{-1}\), kJ\(\cdot\)km\(^{-1}\) and sRPE\(\cdot\)km\(^{-1}\)), according to previous research.\textsuperscript{28,76}

**Statistical analysis**

Before analyzing the assumptions of normality were verified by using Shapiro-Wilk W test and by visual inspection of Q-Q plots. Training intensity and load variables were compared with each other using a multilevel random intercept model using Tukey’s method for pairwise comparisons in R (R: a language and environment for statistical computing; Vienna, Austria). Random effect variability was modeled using a random intercept for each individual participant. Level of significance was established at \(p < .05\). Magnitude-based inferences were used to further evaluate and describe the magnitude of the effects observed. Standardized effect size is reported as Cohen’s \(d\), using the pooled SD as the denominator. Qualitative interpretation of Cohen’s \(d\) was based on the guidelines provided by Hopkins: 0-0.19 trivial; 0.20 – 0.59 small; 0.6 – 1.1.9 moderate; 1.20- 1.99 large; ≥ 2.00 very large.\textsuperscript{95}

**4.3 Results**

A total of 9822 training files were collected and analysed, 7319 training files (74.5%) from male cyclists and 2503 training files (25.5%) from female cyclists. All collected training files contained PO data, were 3868 (52%) and 2049 (82%)
of the training files contained RPE scores for men and women respectively. Furthermore, 4974 (68%) and 1696 (68%) training files contained HR for men’s and women’s respectively.

Descriptive values for both men’s and women’s training data are presented in Table 4.1. Largely higher values \((d = 1.98)\) were observed for mean PO and moderate \((d = 0.65 - 0.82)\) higher values for men were observed for distance and mean relative PO \((\text{W} \cdot \text{kg}^{-1})\). However, women had moderate higher values for mean HR \((d = 0.80)\) and \(\%\text{HR}_{\text{max}}\) \((d = 0.67)\). Small higher values \((d = 0.44-0.57)\) were observed in men’s training data for duration and IF, while no differences were observed for mean RPE-score.

Table 4.1: Volume and intensity training descriptors of men’s and women’s professional cyclists.

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
<th>Differences men vs women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Cohen’s d</td>
</tr>
<tr>
<td><strong>Distance (km)</strong></td>
<td>91.9 (48.1) *</td>
<td>64.1 (38.1)</td>
<td>0.65 – Moderate</td>
</tr>
<tr>
<td><strong>Duration (min)</strong></td>
<td>182 (96) *</td>
<td>145 (72)</td>
<td>0.44 – Small</td>
</tr>
<tr>
<td><strong>Mean PO (W)</strong></td>
<td>191 (32) *</td>
<td>138 (22)</td>
<td>1.98 – Large</td>
</tr>
<tr>
<td><strong>Mean PO (W∙kg(^{-1}))</strong></td>
<td>2.64 (0.45) *</td>
<td>2.30 (0.37)</td>
<td>0.82 – Moderate</td>
</tr>
<tr>
<td><strong>Intensity Factor™</strong></td>
<td>0.59 (0.09) *</td>
<td>0.66 (0.12)</td>
<td>0.57 – Small</td>
</tr>
<tr>
<td><strong>Mean HR (beats∙min(^{-1}))</strong></td>
<td>125 (12)</td>
<td>136 (16)</td>
<td>0.80 – Moderate</td>
</tr>
<tr>
<td><strong>Mean HR (%\text{HR}_{\text{max}})</strong></td>
<td>65.5 (6.3) *</td>
<td>69.8 (6.7)</td>
<td>0.67 – Moderate</td>
</tr>
<tr>
<td><strong>Mean RPE (AU) (6-20 scale)</strong></td>
<td>12.2 (2.3)</td>
<td>12.1 (2.2)</td>
<td>0.02 – Trivial</td>
</tr>
</tbody>
</table>

Abbreviations: PO, power output; HR, heart rate; \(\text{HR}_{\text{max}}\), maximal heart rate; RPE, rating of perceived exertion. Note: Qualitative interpretation of d was based on the guidelines provided by Hopkins et al.\(^{95}\) *Significant difference \((P < 0.05)\).

Figure 4.1 graphically displays the differences between men and women and the distribution of training sessions for duration, TSS, sRPE, TRIMP, RPE and mean HR bands. Table 4.2 presents the absolute and relative load (expressed per kilometre) for both men’s and women’s training data. Absolute and relative total work expressed in kJ spent were moderate to largely higher \((d = 0.99-1.26)\) in men’s training compared to women’s training. Further, relative load expressed as TSS and TRIMP were moderately higher \((d = 0.68-0.76)\) in women’s training compared to men’s training.
Table 4.2. Absolute load and load metrics expressed relative to distance in training sessions of professional male and female cyclists.

<table>
<thead>
<tr>
<th>Men</th>
<th>Women</th>
<th>Differences men vs women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Cohen’s d</td>
</tr>
<tr>
<td><strong>Total work (kJ)</strong></td>
<td>2151 (1238) *</td>
<td>1223 (645)</td>
</tr>
<tr>
<td><strong>TSS (AU)</strong></td>
<td>115.5 (75.2)</td>
<td>111 (68)</td>
</tr>
<tr>
<td><strong>TRIMP (AU)</strong></td>
<td>453 (262.9)</td>
<td>425 (221)</td>
</tr>
<tr>
<td><strong>sRPE (AU)</strong></td>
<td>2495 (1436) *</td>
<td>1850 (1069)</td>
</tr>
<tr>
<td><strong>kJ spent · km⁻¹</strong></td>
<td>23.6 (6.9) *</td>
<td>17.7 (1.3)</td>
</tr>
<tr>
<td><strong>TSS · km⁻¹</strong></td>
<td>1.23 (0.47) *</td>
<td>1.56 (0.49)</td>
</tr>
<tr>
<td><strong>TRIMP · km⁻¹</strong></td>
<td>4.58 (1.75) *</td>
<td>5.72 (1.23)</td>
</tr>
<tr>
<td><strong>sRPE · km⁻¹</strong></td>
<td>25.04 (5.52)</td>
<td>25.91 (4.98)</td>
</tr>
</tbody>
</table>

Abbreviations: TSS, Training Stress Score; TRIMP, Edwards’ training impulse; sRPE, session rating of perceived exertion. Note: Qualitative interpretation of d was based on the guidelines provided by Hopkins et al.95 *Significant difference (P < 0.05).

Figure 4.1: Distribution of training sessions for Duration (A), eTRIMP (B), sRPE (C), TSS (D), RPE (E) and Mean HR (%HRmax) (F) in professional men and women cyclists. eTRIMP indicates Edwards TRIMP, sRPE indicates session RPE, TSS indicates Trainings Stress Score. by Hopkins et al.95 *moderate differences (d≥0.60).
Figure 4.2 graphically displays the differences between men’s and women’s absolute and relative time spent in the 5 different intensity zones for HR and PO. Time spent in both HR and PO low-intensity (i.e. zone 1 and zone 2) was small to moderate higher \((d = 0.42-0.75)\) in men’s training compared to women’s training. Time spent in high intensity HR and PO zones (i.e. zones 4 and 5) was small higher \((d = 0.25-0.42)\) in women compared with men. %Time spent in both HR and PO intensity zone 1 was small higher \((d = 0.39-0.58)\) in men’s training compared to women’s training. %Time spent in high intensity HR and PO zones (i.e. zones 4 and 5) was small to moderate \((d = 0.36-0.64)\) higher in women’s training compared with men’s training.

**Figure 4.2**: Intensity distribution as (A) absolute and (B) %time spent in different HR zones and intensity distribution as (C) absolute and (D) %time spent in PO zones in men’s versus women’s professional cyclists training sessions. ° Individual data-points. HR indicates heart rate; PO indicates power output. Note: Qualitative interpretation of d was based on the guidelines provided by Hopkins et al. \(^95\) * moderates differences \((d>0.60)\).

Figure 4.3 presents the %time spent during training for relative PO bands \((\text{W} \cdot \text{kg}^{-1})\). Women spent a moderately higher %time in the range between 1.51-2.25 \(\text{W} \cdot \text{kg}^{-1}\) \((d = 0.93)\) and men spent moderately higher %time during training in the power bands 3.01-3.75 \(\text{W} \cdot \text{kg}^{-1}\) \((d = 0.84)\) and 3.76-4.50 \(\text{W} \cdot \text{kg}^{-1}\) \((d = 0.85)\).
This study provides a retrospective analysis of a large data-base (~10,000 training sessions) describing the training intensity distribution and training load of professional cyclists and highlights the differences between men and women. On average, men's training sessions were higher in distance, duration, work (kJ spent), absolute PO and relative PO. However, training load (i.e. TSS and TRIMP) of men's and women's training seemed to be similar. In addition, when load values were expressed relative to distance, women show moderate higher values in training compared to men. Similar to the intensity distribution in races, women spent less time training in the low-intensity zones and more in the high-intensity zones compared to men. These descriptive results contribute to a better understanding of the training programs in professional cyclists and the specific differences between men's and women's in training.

Distance, duration and total work (kJ spent) were higher in men's training compared to women's training, which is in line with their respective differences in race demands. Given that total work (kJ spent) is directly related to absolute PO produced, it is no surprise that this was higher in men's training, given their large higher mean PO during training (i.e. 191 [31 W] vs 138 [22 W]). When PO during training is expressed relatively to body weight the differences between male and female cyclists decreased, which is in line with previous research. Inherent physiological differences most likely contribute to the differences in absolute mean PO during training. For example, inherent

Figure 4.3: Power output distribution as %time spent in different power bands, relative to body weight. Note: Qualitative interpretation of d was based on the guidelines provided by Hopkins et al. *moderate differences (d≥0.60).
physiological differences between men and women are related to maximal oxygen uptake and body composition, which results in a higher aerobic capacity observed in professional male cyclists compared with professional female cyclists. Nevertheless, even though overall training volume is lower in women’s training, they seem to compensate this lower volume of training with a higher training intensity. For example, intensity distribution (Figures 3.2 and 3.3) shows that professional women cyclists train at a higher intensity compared to professional men cyclists. Absolute and relatively, men spent more time in the low-intensity (HR and PO) zones (i.e. zones 1 & 2) compared to women, while women spent more time in high-intensity (HR and PO) zones (i.e. zones 4 & 5). This is in line with the higher intensity observed in women’s professional cycling races compared to men’s races. However, the differences in training are smaller compared to the differences in races. For example, women spent ~25% more time in high-intensity zones (HR) during races compared to men in contrast to training were the difference is only ~13% between women and men.

Based on the results reported in this study and the previous results evaluating race demands in men and women’s professional cycling races, it seems that the training of female professional cyclists is adapted (i.e. higher intensity in training compared to men’s training) to their respective race demands (i.e. higher intensity in racing compared to men’s races). Given the descriptive nature of this and the previous study, factors contributing to this remain largely speculative. One factor contributing to this may be that coaches consciously adapt the women’s training program to meet the higher intensity race demands. Another reason for the lower training volume (i.e. hours of training per week) for female professional cyclists, may be the UCI regulations on salaries for Pro Continental and World Tour male cyclists. For male cyclists at that level it is obligatory to receive a certain minimum salary, while in contrast, for females at that level this is not (yet) the case. As a result, some professional female cyclists may be working part time or studying next to their athletic career. Therefore, it could be that a certain number of professional female cyclists can simply not afford to train the same training volume as men. Furthermore, a successful training program requires a balance between overload and recovery, not being able to live like a full-time professional could interfere with recovery, with a lower overall training volume as result. Given
their lower training volume, female professional cyclists might automatically compensate this by increasing their training intensity. Also, differences in race program commitments between male and female professional cyclists may contribute to the observed training differences too. A typical season for a male cyclists contains 70 to 90 race-days with a maximum up to 100 race-days, while a typically women’s season contains 40 to 55 race-days. Therefore, a higher percentage of the training sessions in a male training program would consist of low intensity and recovery riders as it is common to plan an easy or a recovery training after a race. This might contribute to the increased time spent at lower intensity zones in men’s training compared to the women’s training.

Similar to previous research, in race load measurements (TRIMP, kJ spent and sRPE) were ~60% higher compared to training while load measured by TSS is ~120% higher in races compared to training. One of the reasons for the difference between TSS and the other load measurement may be the difference in weighting exercise intensity integrated with the calculation of TSS compared to other measures. While distance, duration and total work (kJ spent) differ between men and women, load values integrating some information relating to the individual characteristic (i.e. TRIMP (max HR), TSS (FTP)) were similar between men and women. This is in line with previous research which found small differences in load (i.e. TRIMP and TSS) between men and women’s races. In addition, when those values were expressed relative to distance (TRIMP∙km⁻¹, TSS∙km⁻¹) women’s have on average a moderate higher load during training sessions compared to men. Which is in line with the reported higher intensity demands of women’s races compared to men’s races. Although differences reported between men and women’s races were large to very large, the differences in training between men and women were only moderate.

Similar to previous research it is important to note, that despite differences in load and intensity distribution between men’s and women’s training, both have the same RPE-scores 12.2 (men) vs 12.1 (women) during training sessions. This is indicating that despite different objective (HR and PO) measurements, subjective perceived exertion can be similar.

A right shift in power bands can be seen when comparing the powerbands of the men with the women in training (Figure 4.3). Women spent
a higher %time in the lower PO bands 0.76 to 3.00 W·kg\(^{-1}\), were men spent a higher %time in PO bands 3.01 to 6.00 W·kg\(^{-1}\). This is probably caused by the differences in maximal PO between men and women. Compared to the power bands reported in races for both men and women, the distribution in training was more centered around 1.50 to 3.75 W·kg\(^{-1}\). Men spent 60% and women 66% of their training between 1.50 to 3.75 W·kg\(^{-1}\) compared to 30% and 37% in races. This is probably caused by the more stochastic PO in races as a result of different levels of races, level of athletes,\(^{42}\) race profile\(^{35,39}\) and race tactics\(^{42}\) and therefore the PO in races is less controlled compared to training. Similar to previous research\(^{28,42}\) a large proportion (15%) of time is spent at PO < 0.75 W·kg\(^{-1}\) due to non-pedaling activity although the reported values in training sessions were slightly lower compared to values in races (20%\(^{42}\) to 25%\(^{28}\)).

**Limitations**

This was a retrospective study were ~10 000 training sessions of 30 professional cyclists (10 female and 20 male) were analyzed. The presented data in this study was primarily collected by the team to monitor the cyclists and secondly for the purpose of this study. This retrospective approached causes some limitations that need to be taken in to account when interpreting the results. A detailed overview of those limitations are already described elsewhere.\(^{28}\) We acknowledge the limitations caused by the retrospective nature of this study. However, this approach made it possible to analyze a large data-set of training data from professional male and female cyclists, which has currently not been published before. This study reports a detailed overview of load and intensity characteristic of professional cyclists and highlights the differences between men’s and women’s training.

**Practical applications**

The descriptive results of this study contribute to a better understanding of the load and intensity characteristics of male and female professional cyclists training sessions. Previous research indicates that training strategies between men and women elite cyclists may not be used interchangeably due to the different demands in races. The results of the presented study show that women’s training load and time spent in high-intensity zones is higher.
compared to men’s training sessions. Although differences in races between men and women were larger compared to the differences in training sessions.

**Conclusions**
To conclude, overall volume and some load parameters (i.e. distance, duration, sRPE, kJ spent and PO) were higher in men’s training compared to women’s training. However, load expressed as TSS and TRIMP were similar in men’s training and women’s training. Furthermore, relative load (TRIMP·km⁻¹, TSS·km⁻¹) and time spent in high-intensity zones was higher during women’s training compared to men’s training, despite similar RPE scores. It seems that women’s training programs compensate their lower training volume, in comparison to men’s training, with a higher training intensity.
Chapter 5

Relationship Between Various Training Load Measures in Elite Cyclists during Training, Road Races and Time Trials.

Teun van Erp • Carl Foster • Jos J. de Koning

Abstract

Purpose: The relationship between various Training Load (TL) measures in professional cycling is not well explored. This study investigates the relationship between mechanical energy spent (in kiloJoules), session Rating of Perceived Exertion (sRPE), Lucia’s TRaining IMPulse (LuTRIMP), and Training Stress Score (TSS) in training, races and Time Trials (TTs).

Methods: For 4 consecutive years, field data were collected from 21 professional cyclists and categorized as being collected in training, racing, or TTs. KiloJoule (kJ) spent, sRPE, LuTRIMP, and TSS were calculated, and the correlations between the various TLs were made.

Results: 11,655 sessions were collected, from which 7,596 sessions had heart rate data and 5,445 sessions had a rating of perceived exertion score available. The r between the various TLs during training was almost perfect. The r between the various TLs during racing was almost perfect or very large. The r between the various TLs during TTs was almost perfect or very large. For all relationships between TSS and one of the other measurements of TL (kJ spent, sRPE, and LuTRIMP), a significant different slope was found.

Conclusion: kJ spent, sRPE, LuTRIMP, and TSS all have a large or almost perfect relationship with each other during training, racing, and TTs, but during racing, both sRPE and LuTRIMP have a weaker relationship with kJ spent and TSS. Furthermore, the significant different slope of TSS versus the other measurements of TL during training and racing has the effect that TSS collected in training and road races differ by 120%, whereas the other measurements of TL (kJ spent, sRPE, and LuTRIMP) differ by only 73%, 67%, and 68%, respectively.
5.1 Introduction

Training Load (TL) is one of the most important parameters for sport scientists and coaches to monitor in elite athletes. TL needs to be high enough to induce a stimulus for adaptation, with evidence that the magnitude of performance adaptation is proportional to the training load. However, too high values for chronic TL are associated with overtraining syndrome and other evidence of mal-adaptation to training. Large increases in TL are also associated with injuries. It seems necessary to increase TL slowly for optimal improvement in elite athletes. Further, to achieve a peak performance at a specific moment, there needs to be an appropriate balance between the increase in fitness and the accumulation of fatigue related to training. This balance is addressed in the rich literature on tapering.

Training programs are mostly based on measurements of external load, which is the TL independent of the individual response to training of the athlete. Coaches can prescribe a training session based on sport specific parameters e.g. Power Output (PO), distance completed, number of throws or duration. External TL is based on the work completed independently of the physiological or perceptual response within the individual athlete. However, the relative physiological stress imposed on the athlete, the internal TL, is a more important determinant of the stimulus for training adaptation than the external TL. Training programs based on measurements of internal TL have mostly used Heart Rate (HR) or the Rating of Perceived Exertion (RPE) to quantify TL.

HR is frequently used to determine TL because the technology is widely available, non-invasive and inexpensive. The use of HR monitoring during exercise is based on the nearly linear relationship between HR and oxygen consumption during steady-state submaximal exercise. There are different means of calculating the TL using HR. Banister et al. introduced the concept of the TRIMP as a marker of TL based on the intensity of the exercise as calculated by the product of the average %HRreserve and the duration of exercise. Multiple studies have validated TL based on HR by correlating TL with other measurements of TL or with one of the variations developed for TRIMP. One of the variations of TRIMP is LuTRIMP, which has been widely used in professional road cycling. LuTRIMP is a summated score based on
the duration spent in each of the three HR zones, weighted for the intensity of the HR zone. Although TL based on HR is frequently used in different sports, many factors may influence the relationship between TL and HR.\textsuperscript{110} The day-to-day variation in HR is approximately $\sim 6$ beats/min\textsuperscript{111} or $< 6.5\%$\textsuperscript{112} when factors influencing HR such as hydration, environment and fatigue are not controlled.\textsuperscript{110} This can make measuring internal TL based on HR challenging in the field. Further, the TRIMP method relies on the availability of known values of resting and maximal HR.

Foster et al.\textsuperscript{17} introduced a non-invasive, inexpensive and easy to use method to determine internal TL, based on a simple modification of the well-regarded RPE scale.\textsuperscript{18} The TL of the exercise session is defined by multiplying the RPE of the entire session by the duration of the session in minutes; the session RPE (sRPE). This method made the measurement of TL independent of equipment. sRPE has been shown to be a valid and reliable measure of TL\textsuperscript{113} and has a good correlation with TRIMP in elite cyclists during different stages races.\textsuperscript{76} sRPE is minimally influenced by the time of measuring after training,\textsuperscript{100} but is influenced by environmental conditions\textsuperscript{114} and hydration status.\textsuperscript{115} Therefore the measurement of TL by the use of sRPE can be challenging in the field as well.

In 1986 Ulrich Schoberer develop the first power meter (SRM) for the outdoor bicycle, bringing a new power-based training approach to elite cycling. Coggan and Allen developed an approach to estimate TL based on continuous collection of the PO of a cyclist using a power meter normalized using an individual threshold parameter (Functional Threshold Power (FTP), called Trainings Stress Score (TSS)).\textsuperscript{23} A workout of one hour on FTP represents a TSS-score of 100 AU. Sanders et al.\textsuperscript{36} investigated the effect of fatigue on the relationship between various TLs in elite cyclists. However, to the best of our knowledge, no study has investigated the effect of different situations on the relationship between various TLs. Therefore, the aim of this study was to investigate the relationship between various TL measures based on HR, sRPE and PO during training, road racing, and time trials in a group of elite cyclists, across multiple seasons. We hypothesized that the relationship between TLs would be highest during training and would be lower during road races and time trials, because, as mentioned previously, factors that influence internal TL are harder to control in competition.
5.2 Methods

Design
During 4 consecutive years we collected field data within a professional cycling team (1-year Pro-continental and 3-year World Tour) with the aim of monitoring the cyclists and analyzing their performance. Depending on the athlete’s tenure on the team, the data set of an individual cyclist contains data ranging from 1 to 4 years. If a cyclist was not able to train for a period of 3 months or more, because of illness or an injury, the data set of this particular year was excluded. Institutional ethics approval was granted and, in agreement with the Helsinki Declaration, written informed consent was obtained from the participants.

From as many as possible cycling sessions HR, RPE and PO were collected and uploaded by the cyclists to a central database. All data was categorized as being collected in training, road racing or time trials according to the UCI, based on www.procyclingstats.com. The power-files categorized as racing or time trials could contain data from warm-up, cool-down or reconnaissance and therefore may be longer than the actual race or time-trial. All data sets were visually checked and incomplete data sets were excluded.

Subjects
Twenty-one highly trained professional cyclists participated in this study. During the 4 years of analysis the cyclists won combined, more than 100 UCI races, of which 45 victories were in the World Tour, including 29 grand tour stages. Two cyclists also finished in the top 10 of the general classification of a grand tour. All riders had experience racing on World Tour level and trained or raced ~30 000 to 35 000 km per year. Their average 20-minute Mean Maximal Power (MMP) of 413 (32) W or 5.55 (0.8) W·kg⁻¹ can be seen as a measure of their excellence.

kJ spent
For every training, road race or time trial the total mechanical energy spent (in kJ) was calculated from the PO as measure of external TL. PO was collected with the use of SRM power-meters (SRM, Jülich, Welldorf, Germany) and Pioneer power meters (Pioneer, Kawasaki, Japan). All riders were informed
about the importance of the zero-calibration and were instructed to do the zero-calibration before every ride, however this was not controlled by the team.

**sRPE**

All riders were obligated by the team to fill in their training logbook daily, as soon as possible after the race or training. The RPE was obtained on a 6-20 scale\(^{116,117}\) based on the question in the logbook: ‘How hard was your day?’.\(^{18}\) The sRPE was calculated by multiplying the RPE of the session by the duration of the exercise in minutes (Equation 5.1) according to Foster et al.\(^{10,17,113,118}\) as a measure of internal TL based on RPE. As a general rule, the total time of the training session or race including warm-up and cool-down was included in the duration element of sRPE.

(Equation 5.1)

\[
sRPE = RPE \cdot \text{Duration (min)}
\]

The duration of the exercise was based on the collected data of that particular day. When the power meter recorded speed, the duration of the exercise was based on the time that speed was recorded. When no speed was recorded the duration of the session was based on the time that HR was recorded. In the case that both speed and HR were missing, the duration of the exercise was manually determined. The session was excluded from this study when it was not possible to measure exercise duration.

**LuTRIMP**

Lucia’s TRIMP (LuTRIMP)\(^{2,22}\) was used to calculate the internal TL based on HR. HR was recorded with different brands of heart rate monitors (Suunto, Vantaa, Finland & Pioneer, Kawasaki, Japan). LuTRIMP was calculated by multiplying the time spent in 3 HR-zones (zone 1: below VT, zone 2: between VT and RCT and zone 3: above RCT) by a coefficient relative to each intensity zone (k=1 for zone 1, k=2 for zone 2, k=3 for zone 3) according to Equation 5.2.

(Equation 5.2)

\[
luTRIMP = (\text{Duration (min)in zone 1} \cdot 1) + (\text{Duration (min)in zone 2} \cdot 2) + (\text{Duration (min)in zone 3} \cdot 3)
\]
Originally the three HR-zones where LuTRIMP is based on, are computed based on responses during laboratory exercise testing. However, no laboratory exercise testing was implemented within the team. Therefore we determined the HR-zones by the highest HR recorded during every season (from November until October) and used HR-zones as described by Seiler. All HR data were manually checked for errors and when necessary rejected from analysis.

**TSS**

TSS was used as measure for internal TL based on PO. For every ride TSS was calculated according to Equation 5.3.

\[
TSS = \left( \frac{t \cdot NP \cdot IF}{FTP \cdot 3600} \right) \cdot 100
\]

(Equation 5.3)

Were \( t \) is the duration of the session in seconds and \( IF \) the intensity factor (see Equation 5.5). \( NP \) is the normalized power as calculated with Equation 5.4, where \( P_i \) is the floating mean power during 30 seconds time segment and \( N \) is the total number of time segments. Session duration \( t \) is obtained based on speed, HR or manually as described above. The Functional Threshold Power (FTP) was determined as 95% of the highest 20 minutes MMP obtained in every season.

\[
NP = \sqrt[4]{\frac{1}{N} \sum_{i=1}^{N} p_i^4}
\]

(Equation 5.4)

\[
IF = \left( \frac{NP}{FTP} \right)
\]

(Equation 5.5)
Statistical analysis

Descriptive data are reported as means (standard deviation). MATLAB and Statistics Toolbox (Release 2012b, The MathWorks, Inc., Natick, Massachusetts, United States) were used to analyze the data and for the statistics. Differences between the four data sets used to measure TL were determined by using N-way of Analysis of Variance (ANOVAN). When 10 or more training, races or time trials sessions were collected for an individual cyclist, the relationship between all various TLs was determined by using the Pearson correlation coefficient. Regression lines are force through zero, correlation coefficients (r and slopes) values are described and uncertainties in the correlation coefficients are presented as 95% CI. Further ANOVAN was used to determine significant difference between the slopes of the relationships during training, racing or TTs. As recommended by Hopkins, Fisher’s z-transformation was used to transform correlation coefficients (r) before averaging and calculating 95% confidence intervals (CI). The following criteria were adopted to interpret the magnitude of the correlation(r) between the measures: < 0.1 trivial, 0.1-0.3 small, 0.3-0.5 moderate, 0.5-0.7 large, 0.7-0.9 very large, and 0.9-1.0 almost perfect. The level of statistical significance was set at 0.05.

5.3 Results

PO from 11 655 sessions were collected from which 7596 sessions had HR data available and 5445 sessions had an RPE score. Based on the combined presence of RPE, HR and PO the kJ spent, sRPE, LuTRIMP and TSS were calculated as measures of TL and are presented together in Table 5.1 for all data combined and in Tables 5.2, 5.3 and 5.4 for the training sessions, road races and TTs respectively.
Table 5.1: Number of observations (N) and averages (SD) for session duration (time), power output (power), external TL and internal TL for RPE (sRPE), HR (luTRIMP) and power output (TSS) measures of TL of all exercise sessions.

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Time (min)</th>
<th>Power (W)</th>
<th>TL (AU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>kJ spent</td>
<td>11655</td>
<td>207</td>
<td>200</td>
<td>2572</td>
</tr>
<tr>
<td>sRPE</td>
<td>5455</td>
<td>211</td>
<td>198&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2889</td>
</tr>
<tr>
<td>luTRIMP</td>
<td>7596</td>
<td>216&lt;sup&gt;a&lt;/sup&gt;</td>
<td>201</td>
<td>253</td>
</tr>
<tr>
<td>TSS</td>
<td>11655</td>
<td>207</td>
<td>200</td>
<td>153</td>
</tr>
</tbody>
</table>

Abbreviations: HR, heart rate; kJ, kilojoules; LuTRIMP, Lucia’s training impulse; sRPE, session rating of perceived exertion; TL, training load; TSS, training stress score.

<sup>a</sup>Significantly different from the 3 other data sets.

Table 5.2: Number of Observations (N) and Averages (SD) for Session Duration (Time), Power Output (Power), External TL (kJ spent), and TL for RPE (sRPE), HR (luTRIMP) and Power Output (TSS) Measures of TL of Training Sessions.

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Time (min)</th>
<th>Power (W)</th>
<th>TL (AU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>kJ spent</td>
<td>7867 (67%)</td>
<td>180&lt;sup&gt;a&lt;/sup&gt; (98)</td>
<td>190&lt;sup&gt;b&lt;/sup&gt; (36)</td>
<td>2137 (1261)</td>
</tr>
<tr>
<td>sRPE</td>
<td>4084 (75%)</td>
<td>194 (96)</td>
<td>192 (32)</td>
<td>2476 (1460)</td>
</tr>
<tr>
<td>luTRIMP</td>
<td>5400 (71%)</td>
<td>194 (97)</td>
<td>193&lt;sup&gt;b&lt;/sup&gt; (29)</td>
<td>212 (113)</td>
</tr>
<tr>
<td>TSS</td>
<td>7867 (67%)</td>
<td>180&lt;sup&gt;a&lt;/sup&gt; (98)</td>
<td>190&lt;sup&gt;b&lt;/sup&gt; (36)</td>
<td>114 (77)</td>
</tr>
</tbody>
</table>

Abbreviations: HR, heart rate; kJ, kilojoules; LuTRIMP, Lucia’s training impulse; RPE, rating of perceived exertion; sRPE, session rating of perceived exertion; TL, training load; TSS, training stress score.

<sup>a</sup>Significantly different from the 2 other data sets.

<sup>b</sup>Significantly different from luTRIMP.

Table 5.3: Number of Observations (N) and Averages (SD) for Session Duration (Time), Power Output (Power), External TL (kJ spent), and TL for RPE (sRPE), HR (luTRIMP) and Power Output (TSS) Measures of TL of Road Races.

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Time (min)</th>
<th>Power (W)</th>
<th>TL (AU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>kJ spent</td>
<td>3457 (30%)</td>
<td>281 (61)</td>
<td>218 (36)</td>
<td>3701 (988)</td>
</tr>
<tr>
<td>sRPE</td>
<td>1241 (23%)</td>
<td>279 (62)</td>
<td>217 (33)</td>
<td>4137 (1217)</td>
</tr>
<tr>
<td>luTRIMP</td>
<td>2056 (27%)</td>
<td>281 (61)</td>
<td>219 (33)</td>
<td>356 (92)</td>
</tr>
<tr>
<td>TSS</td>
<td>3457 (30%)</td>
<td>281 (61)</td>
<td>218 (36)</td>
<td>251 (72)</td>
</tr>
</tbody>
</table>

Abbreviations: HR, heart rate; kJ, kilojoules; LuTRIMP, Lucia’s training impulse; RPE, rating of perceived exertion; sRPE, session rating of perceived exertion; TL, training load; TSS, training stress score.
Relationship between various training load measures.

For the training sessions *almost perfect* correlations were found between all relationships analyzed regardless of where the TL was based on (Tables 5.5-5.10). For road racing the relationship between kJ spent vs TSS and sRPE vs kJ spent were found to be *almost perfect* where all other relationships in races are *very large*. For the TTs an *almost perfect* correlation was found between LuTRIMP vs kJ spent, TSS vs kJ spent and TSS vs LuTRIMP, where a *very large* relationship was found between sRPE vs kJ spent, TSS vs sRPE and sRPE vs LuTRIMP. Slopes did significantly differ between training vs races and training vs TTs for all relationship were TSS was involved (TSS vs kJ spent, TSS vs sRPE and TSS vs LuTRIMP). Where slopes did not differ between training, race and TTs session for the relationships based on the other TLs (sRPE vs kJ spent, LuTRIMP vs kJ spent and sRPE vs LuTRIMP). See Figure 5.1 for an example of a data-set of one rider.

**Table 5.4:** Number of Observations (N) and Averages (SD) for Session Duration (Time), Power Output (Power), External TL (kJ spent), and TL for RPE (sRPE), HR (LuTRIMP) and Power Output (TSS) Measures of TL of Time Trials.

<table>
<thead>
<tr>
<th>Time Trials</th>
<th>N</th>
<th>Time (min)</th>
<th>Power (W)</th>
<th>TL (AU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>kJ spent</td>
<td>331</td>
<td>81 (51)</td>
<td>249 (87)</td>
<td>1411 (659)</td>
</tr>
<tr>
<td>sRPE</td>
<td>120</td>
<td>85 (48)</td>
<td>235 (77)</td>
<td>1298 (723)</td>
</tr>
<tr>
<td>luTRIMP</td>
<td>140</td>
<td>80 (52)</td>
<td>255 (78)</td>
<td>116 (65)</td>
</tr>
<tr>
<td>TSS</td>
<td>331</td>
<td>81 (51)</td>
<td>249 (87)</td>
<td>82 (45)</td>
</tr>
</tbody>
</table>

Abbreviations: HR, heart rate; kJ, kilojoules; LuTRIMP, Lucia’s training impulse; RPE, rating of perceived exertion; sRPE, session rating of perceived exertion; TL, training load; TSS, training stress score.

**Table 5.5:** Average Correlation Coefficient and the Regression Coefficient (Slope) Between sRPE and kJ spent for Training Sessions, Road Racing and Time Trials Presented with 95% CIs.

<table>
<thead>
<tr>
<th>sRPE versus kJ spent</th>
<th>N cyclists</th>
<th>N datasets</th>
<th>r (95% CI)</th>
<th>Slope (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>21</td>
<td>4084</td>
<td>0.97 (0.96-0.97)</td>
<td>0.90 (0.87-0.94)</td>
</tr>
<tr>
<td>Road racing</td>
<td>19</td>
<td>1241</td>
<td>0.91 (0.89-0.94)</td>
<td>0.85 (0.80-0.89)</td>
</tr>
<tr>
<td>Time Trials</td>
<td>4</td>
<td>60</td>
<td>0.89 (0.67-0.97)</td>
<td>0.90 (0.74-1.06)</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; kJ, kilojoules; sRPE, session rating of perceived exertion.
### Table 5.6: Average Correlation Coefficient and the Regression Coefficient (Slope) Between LuTRIMP and kJ spent for Training Sessions, Road Racing and Time Trials presented with 95% CIs.

<table>
<thead>
<tr>
<th></th>
<th>N cyclists</th>
<th>N datasets</th>
<th>r (95% CI)</th>
<th>Slope (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Training</strong></td>
<td>21</td>
<td>4084</td>
<td>0.97 (0.96-0.97)</td>
<td>10.8 (10.5-11.1)</td>
</tr>
<tr>
<td><strong>Road Racing</strong></td>
<td>19</td>
<td>1241</td>
<td>0.85 (0.82-0.87)</td>
<td>10.1 (9.6-10.7)</td>
</tr>
<tr>
<td><strong>Time Trials</strong></td>
<td>4</td>
<td>60</td>
<td>0.93 (0.90-0.94)</td>
<td>9.8 (9.0-10.5)</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; kJ, kilojoules; LuTRIMP, Lucia’s training impulse

### Table 5.7: Average Correlation Coefficient and the Regression Coefficient (Slope) Between TSS and kJ spent for Training Sessions, Road Racing and Time Trials Presented with 95% CIs.

<table>
<thead>
<tr>
<th></th>
<th>N cyclists</th>
<th>N datasets</th>
<th>r (95% CI)</th>
<th>Slope (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Training</strong></td>
<td>21</td>
<td>4084</td>
<td>0.96 (0.95-0.96)</td>
<td>17.9 (17.2-18.7)</td>
</tr>
<tr>
<td><strong>Road Racing</strong></td>
<td>19</td>
<td>1241</td>
<td>0.94 (0.94-0.95)</td>
<td>14.7 (14.1-15.3)</td>
</tr>
<tr>
<td><strong>Time Trials</strong></td>
<td>4</td>
<td>60</td>
<td>0.98 (0.97-0.99)</td>
<td>13.7 (13.2-14.3)</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; kJ, kilojoules; TSS, training stress score. *Significantly different form the 2 other data sets.

### Table 5.8: Average Correlation Coefficient and the Regression Coefficient (Slope) Between TSS versus sRPE for Training Sessions, Road Racing and Time Trials Presented with 95% CIs.

<table>
<thead>
<tr>
<th></th>
<th>N cyclists</th>
<th>N datasets</th>
<th>r (95% CI)</th>
<th>Slope (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Training</strong></td>
<td>21</td>
<td>4084</td>
<td>0.95 (0.95-0.96)</td>
<td>20.3 (19.3-21.3)</td>
</tr>
<tr>
<td><strong>Road Racing</strong></td>
<td>19</td>
<td>1241</td>
<td>0.86 (0.82-0.89)</td>
<td>17.4 (16.6-18.3)</td>
</tr>
<tr>
<td><strong>Time Trials</strong></td>
<td>4</td>
<td>60</td>
<td>0.86 (0.57-0.96)</td>
<td>15.3 (13.8-16.7)</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; TSS, training stress score; sRPE, session rating of perceived exertion. *Significantly different form the 2 other data sets.
This is the first study that has analyzed internal and external TL based on sRPE, HR and PO in professional cycling, for multiple riders, across multiple seasons, in three different types of cycling exercise. Some studies have presented internal TL based on HR (LuTRIMP) and RPE data (sRPE) during grand tours\(^22,69\) or stage races of between 5 and 7 days in durations\(^76\) and TTs\(^119\) but none have reported LuTRIMP and sRPE during training. Further TSS has only been investigated during training in well-trained cyclists\(^16\) and never during road races. This study presents various measures of TL (kJ spent, sRPE, LuTRIMP and TSS) during training, competition and TTs in professional cycling in relation with each other. All four measures of TL have \textit{very large or almost perfect} correlations with each other measured in training, road racing or time trials. However, during road racing the correlation seems to be smaller when
the relationship is based on either LuTRIMP or sRPE, whereas the correlation with kJ spent and TSS for competition stays within the 95% confidence intervals.

Another finding of this study is that the slopes between the relationship of various TLs (sRPE, LuTRIMP and kJ spent) did not differ between measurements collected during training, racing or TT. While the slopes differ between training vs racing and training vs TTs when the relationship was based on TSS. In general, all relationships between various TL measures reported very large or almost perfect relationships with each other, this could be caused by the fact that all measurements are based on time. So, despite that TLs are also based on measures of intensity, RPE, HR and PO, they all have time as a common parameter. The sessions used in this study are from highly trained endurance athletes and therefore it could be possible that the long duration of the sessions overwhelms the effect of RPE, HR and PO in the measurement of TL.

We expected that TSS would have an almost perfect correlation with kJ spent during training, road racing and time trials because TSS is calculated with the use PO and duration which is also the case for kJ Spent. The biggest difference between TSS and kJ spent is that FTP is used to normalize for differences in absolute PO between different cyclists. In this study we determined the FTP for each cyclist by using the 20-minute MMP of the particular season. Nimmerrichter et al. showed that 20-minute MMP is changing by 0.4 W·kg⁻¹ during the season and therefore that taking the highest 20-minute MMP to calculate FTP for an entire season can cause an overestimation of TSS in the build-up periods of the year compared to the part of the season when the highest 20 minute MMP was measured.

We observed a significant shorter average training duration for the TSS observations (Table 5.2) and therefore the LuTRIMP and sRPE reported for training could be slightly overestimated compared to TSS. The shorter duration of the TSS observations found its origin in that in practice the importance of always recording HR and RPE during shorter recovery rides is seen as low, while PO was automatically recorded. This could give a slightly overrepresentation of these shorter rides in the TSS data.
Figure continued on the next page
Figure 5.1: Relationships between various training load measures (kJ spent, LuTRIMP, sRPE and TSS) for a representative subject collected during training (A), road races (B) and time trials (C). Each dot represents one training or one competitive event. Abbreviation: LuTRIMP, Lucia’s training impulse; sRPE, session rating of perceived exertion; TSS, training stress score.
For all the relationships where TSS is involved we found a significantly different slope between training & racing and training & TTs. Therefore, the collection of TSS during training versus race and TTs is different compared to the other measurement of TL (kJ spent, LuTRIMP and sRPE). This is clearly evident when looking to the difference of the average TL during training and racing. When TL is based on TSS, on average the TL in races was 120% higher compared to training. This increase in TL is larger than the increase in TL of the other measurements which is 73%, 67%, and 68% for kJ spent, LuTRIMP and sRPE respectively. Around 10% of this difference can be explained by the significantly lower duration and PO (Table 5.2) measured in the TSS sessions. It seems likely that despite the good relationship between TSS and various TLs, there is a difference between TSS in road races and training, probably caused by the importance of IF in the calculations of TSS. This should be clarified in further analysis.

In line with others, we reported high correlations between various measurements of TL and LuTRIMP. We found an almost perfect relationship between LuTRIMP vs kJ spent, LuTRIMP vs TSS and LuTRIMP vs sRPE during training. However, the present results reveal a weaker relationship in road races between LuTRIMP vs kJ spent, LuTRIMP vs TSS and LuTRIMP vs sRPE compared to the relationship in training. This was expected because during racing the factors influencing HR are harder to control. Multiple factors could result in a lower or higher HR compared to PO or RPE in races, with the result a weaker relationship with kJ spent and TSS. It could be that fatigue plays a major role in the slightly weaker relationship, because in professional cycling over 75% of the races are stage-races. This means that cyclists are racing for multiple (up to 10) days without recovery days, with an increase in kJ spent of 73% compared to training (Tables 5.2 and 5.3). Halson showed a reduction of 9.3% in maximum HR after 1 week of intensified training with limited recovery. Further Lucia et al. found an average decline of maximum HR of 0.389/day during the Vuelta and 0.351/day during the Tour de France, which was confirmed by Rodriguez-Marroyo et al. Because the HR-zones to calculate LuTRIMP are not modified for this suppression in HR, the LuTRIMP scores are reduced in relation to the external TL during multiple-day races. Thus, LuTRIMP could provide trainers and cyclists with non-accurate TL in periods with a continuous high volume of external TL. Recently, Rodriguez-Marroyo et al.
have recalculated the TRIMP, based on exercise tests performed immediately before and immediately after grand tours. They have shown that if the anchors for the zones in the TRIMP calculation are adjusted for the declining maximum HR during the course of a grand tour, the percentage of time within various zones remains relatively constant from week to week. Thus, by taking the suppression of HR into account when data are collected during a multiple stage race, a better estimation of internal TL on the basis of HR can be made. A higher HR compared to PO could occur when riders are well-rested after a taper period before the most important races and this has the opposite effect of fatigue on the HR-PO relationship. Further during races obtaining nutrition or fluids in a timely way is more difficult compared to training. Because race stress makes timing harder and the availability of nutrition or fluids during races is limited to supplies from the car or feeding zone. A limited availability of nutrition and fluids could result in higher cardiac drift. Furthermore riding in a hot environment results in a higher cardiac stress. Races start at a fixed schedule therefore it is impossible to avoid the hottest moments during races in contrast of training where riders can select their own timeslot and avoid those hottest moments. All mentioned factors will result in a higher HR compared to PO and will therefore result in a higher LuTRIMP compared to kJ spent or TSS.

We found similar values for LuTRIMP during road racing (356 AU) compared to the values reported by Lucia (2003) during the Tour de France (375 AU) and Vuelta a España (362 AU). Despite the fact that we used maximum HR to determine the HR-zones instead of a laboratory incremental cycling test, our values for LuTRIMP seems to be comparable to previous studies.

Despite many factors that can influence sRPE, such as hydration status and environmental conditions, we found an almost perfect relationship between sRPE and various TLs during training and a very large relationship during road races where sRPE vs kJ Spent was almost perfect. The slightly lower correlation between sRPE and various TLs obtained during road races can be caused by the factors mentioned above, because they can be less controlled during races. Further, professional cyclists have a periodization schedule to be in the best shape at the start of important road races. Therefore, it could be that their physiological condition influences the relationship between sRPE and the other measurements of TL. Furthermore, stage races
can take up to 10 days of racing without a rest-day and it could be possible that accumulating physical or mental fatigue makes this relationship weaker during road races. However, there is still a very large relationship between sRPE and other measures of TL during road races. Thus, sRPE seems to be a reliable and practical tool to measure TL. Because the 6-20 “classical” version of the RPE scale seems to translate for European athletes better than the 0-10 Category Ratio version of the scale, we used the classical version, rather than the Category Ratio version used by Foster et al.\textsuperscript{10,17,100,113,118} However, recent evidence from our laboratory\textsuperscript{123} suggests that the pattern of response of the two versions of the RPE scale are very consistent, suggesting that, although leading to numerically distinct sRPE values, the overall impression of the training session is unaffected by the version of RPE scale used.

The results of our study indeed showed smaller correlations during races and time trials for TLs based on internal measurements (RPE and HR). For coaching these uncontrollable influences on internal TL (LuTRIMP & sRPE) are of great interest and for this reason it would be discussable if a perfect correlation is desirable as it means that physiological or mental fatigue are not measured and the additional value of measuring internal TL is absent.

**Limitations**

The reported results are from an observational study over multiple years with 21 professional cyclists. The presented data was primarily collected for monitoring professional cyclists and secondly for the purpose of this study. Therefore, the collected data has some limitations. The measurement of PO was done with 2 different brands of power-meters. Riders had different bicycles equipped with different power-meters and the riders had the responsibly to calibrate the power-meters every day. To our knowledge, a systematic comparison of different power-meters has not been reported in the literature. Further the riders were obligated by the team to provide RPE-scores, but we did not control the timing of obtaining the RPE-scores. Therefore it could be that RPE is recorded later than 30 min after the exercise as recommend by Foster et al.,\textsuperscript{10,17,118} because of unforeseen issues (logistics, bad internet connection, press, podium ceremonies) or forgotten and filled in later. However, recent data have shown that the sRPE is remarkably robust relative to the timing between the exercise bout and the collection of the sRPE.
score. In the studied subjects, some did strength training adjacent to cycling training, so the RPE-score can be influenced by the strength training. Also, HR was measured with different heart rate monitors, although meaningful differences between brands of radio-telemetric heart rate belts have not been reported.

**Conclusions**

TLs based on RPE (sRPE), HR (LuTRIMP) and PO (kJ spent & TSS) are all reliable measurements for measuring TL in training, road racing and TTs in professional cyclists. However, the difference between TSS collected in training and road-races (120%) is unexpectedly higher compared to this difference in the other measures in TL (kJ spent, sRPE and LuTRIMP respectively 73%, 67%, and 68%). This is cause by the significantly different slopes of TSS vs the other TLs during training and races & TTs. In further research the behavior of LuTRIMP, sRPE and TSS during road races needs to be investigated to better understand the relationship with external TL.

**Application to practice**

sRPE and LuTRIMP are methods for measuring internal TL which are simple and low cost compared to TSS and can easily be used when there are no possibilities to train or complete races with a power-meter. However, during races both methods to estimate internal TL have a weaker relationship with kJ spent than the TSS and may therefore be less reliable in comparison with TSS. On the other hand, despite the good correlation with kJ spent, TSS shows unexpected higher values during racing compared to the increase of the other measurements of TL. At the same time, the lack of a true “gold standard” for evaluating the TL, particularly the internal TL which may respond to progressive fatigue, makes the small differences observed within the current data perhaps less than critical. It suggests that any method of monitoring TL which is consistently applied and discussed between coach and athlete may be more or less equivalent in net value.
Chapter 6

The Influence of Exercise Intensity on the Association Between Kilojoules Spent and Various Training Loads in Professional Cycling.

Teun van Erp • Marco Hoozemans • Carl Foster • Jos J. de Koning

*Int J Sports Physiol Perform. [Epub ahead of print]*
Abstract

Purpose: A valid measure for Training Load (TL) is an important tool for cyclists, trainers, coaches and sport scientists involved in professional cycling. The aim of this study is to explore the influence of exercise intensity on the association between kiloJoules (kJ) spent and different measures of TL to arrive at valid measures of TL.

Methods: Four years of field data were collected from 21 cyclists of a professional cycling team, including 11 716 training and race sessions. kJ spent was obtained from power output measurements and other TLs were calculated based on the session Rating of Perceived Exertion (sRPE), heart rate (Lucia TRIMP [luTRIMP]) and power output (Training Stress Score [TSS]). Exercise intensity was expressed by the Intensity Factor (IF). To study the effect of exercise intensity on the association between kJ spent and various other TLs (sRPE, luTRIMP and TSS), data from low- and high-intensity sessions were subjected to regression analyses using Generalized Estimating Equations (GEE).

Results: This study shows that the IF is significantly different for training and race sessions (0.59 [0.03] vs. 0.73 [0.03]). Significant regression coefficients show that kJ spent is a good predictor of sRPE, luTRIMP, as well as TSS. However, IF does not influence the associations between kJ spent and sRPE and luTRIMP, while the association with TSS is different when sessions are done with low or high IF.

Conclusion: It seems that the TSS reacts differently to exercise intensity than sRPE and luTRIMP. A possible explanation could be the quadratic relation between IF and TSS.
6.1 Introduction

Competitive road cycling is one of the sports with the highest volume in Training Load (TL), therefore professional cyclists experience high physiological demands. It is important to have a valid measure of those physiological demands to optimize training outcomes and to prevent the overtraining syndrome, which is associated with a high volume in TL.\textsuperscript{64} Furthermore, spikes in TL are associated with increased risk of injuries.\textsuperscript{65} To achieve a peak performance at a specific moment, an appropriate balance between the increase in fitness and the accumulation of fatigue related to training is needed.\textsuperscript{104,105}

Training programs are mostly based on measurements of external load, which is the TL independent of the individual characteristics of the athlete, for example, Power Output (PO) or distance completed. However, the relative physiological stress imposed on the athlete, the internal TL, is a larger determinant of the stimulus for training adaptation than the external TL.\textsuperscript{107} Two frequently used measures of internal TL are the Rating of Perceived Exertion (RPE) and TRIMP. Foster et al.\textsuperscript{17} measured TL by using RPE\textsuperscript{18} over the entire duration of exercise bouts and multiplying this number by the duration of the exercise session in minutes, the session RPE (sRPE). Banister et al.\textsuperscript{19} introduced the concept of TRIMP as a marker of TL based on the intensity of the exercise, calculated as the product of the average heart rate (HR) reserve and the duration of exercise. There are different variations of TRIMP used in literature. One of them is Lucia’s TRIMP (luTRIMP), which is a summation method based on heart rate zones and is often used in professional road cycling.\textsuperscript{2,16,22}

In 1986, Ulrich Schoberer developed the first power meter for the outdoor bicycle, bringing a new power-based training approach to professional cycling. Allen and Coggan developed an approach to quantify TL based on PO of cyclists using a power meter, the Training Stress Score (TSS).\textsuperscript{23} Almost all training analysis software are using TSS as their most important measurement of TL, resulting in being TSS the measure of TL for many trainers and coaches in professional cycling. Despite the fact that TSS is highly used, there is almost no research available about the validity of TSS as measure for TL. Furthermore, van Erp et al.\textsuperscript{103} showed that TL measured during training
differed from TL measured during races. It was found that during races the kilojoule (kJ) spent was 73% higher, with a comparable increase in luTRIMP, and sRPE (68% and 67% respectively), but with an increase of 120% in TSS.

Coggan and Allen\textsuperscript{23} calculated TSS as:

\begin{equation} \label{eq:6.1}
TSS = \left( \frac{t \times NP \times IF}{FTP \times 3600} \right) \times 100
\end{equation}

where \( s \) is the time in seconds; \( NP \) is the Normalized Power during the ride; \( FTP \) is Functional Threshold Power for the cyclist; and \( IF \) is the Intensity Factor, which is defined as the ratio between \( NP \) and \( FTP \):

\begin{equation} \label{eq:6.2}
IF = \left( \frac{NP}{FTP} \right)
\end{equation}

Since \( IF \) is described by the ratio between \( NP \) and \( FTP \), \( NP \) is one of the most important factors and is described by Equation 6.3:

\begin{equation} \label{eq:6.3}
NP = \sqrt[N]{\frac{1}{N} \sum_{i=1}^{N} p_i^4}
\end{equation}

where \( p_i \) is the floating mean power during time segment \( i \), and \( N \) is the total number of time segments. Kuylenstierna (2016) combined Equation 6.1 and Equation 6.2, to a simplified expression of TSS by Equation 6.4:

\begin{equation} \label{eq:6.4}
TSS = h \times IF^2 \times 100
\end{equation}

where \( h \) is session time in hours. By using \( IF \) rather than absolute power, TSS becomes a metric normalized to the level of each athlete. As shown in Equation 6.4, TSS has a quadratic relationship with exercise intensity. Internal TL can also be quantified by relative VO\textsubscript{2} consumption, which is linearly related to relative heart rate and RPE.\textsuperscript{13,113} Furthermore, other measurements of internal TL (e.g. iTRIMP\textsuperscript{99} and banTRIMP\textsuperscript{105}) have an exponential weighting factor based
on the classically described exponential rise of blood lactate when exercise intensity exceeds the lactate threshold. Therefore, measurements of TL are based on a linear or exponential relationship with exercise intensity in contrast to the quadratic relationship between TSS and exercise intensity. This quadratic relationship between TSS and IF could be the reason for the larger increase in TSS compared to the increase in sRPE and luTRIMP during racing as found by van Erp et al. because racing is probably done with a higher intensity compared to training.

In the present study, we investigated whether racing is done with a higher IF compared with training and how IF (low or high) affects the association between kJ spent and various TLs (sRPE, luTRIMP and TSS). We hypothesize that IF is higher for races than for training and that the association between kJ spent and TSS is different between training and races, while the association between kJ spent and sRPE and luTRIMP are not affected by IF.

6.2 Methods

This study was conducted with the same data set as described by van Erp et al. and therefore only a summary of the data collection is described below. For more details see van Erp et al.

Design

Four years of field data were collected within a professional cycling team with the aim of monitoring the cyclists and analyzing their performance. Institutional ethics approval was granted and, in agreement with the Declaration of Helsinki, written informed consent was obtained from the participants. From as many as possible cycling sessions (training as well as races) HR, RPE and PO data were collected and uploaded by the cyclists to a central database. All data sets were visually checked and incomplete data sets were excluded.

Subjects

A total of 21 highly trained professional cyclists participated in this study. Their average 20-minute Mean Maximal Power (MMP) of 413 (32) W or 5.55 (0.8) W·kg⁻¹ and can be seen as a measure of their excellence.
**kJ Spent**

From every session the total mechanical energy spent (in kJ) was calculated from the PO as measure for external TL. PO was collected with the use of SRM power meters (SRM, Jülich, Welldorf, Germany) and Pioneer power meters (Pioneer, Kawasaki, Japan). All cyclists were informed about the importance of the zero-calibration and were instructed to do the zero-calibration for every ride. The duration of the exercise was based on the collected data of that particular day. The session was excluded from this study when it was not possible to measure exercise duration. Warm up and cool down riding, as well as the main training or racing was included in the calculation of total mechanical energy.

**Session Rating of Perceived Exertion**

The RPE of every session was obtained on a 6-20 scale, based on the question: ‘How hard was your day?’ As a measure of TL based on RPE, the sRPE was calculated by multiplying the RPE of the session by the duration of the exercise in minutes, according to Foster et al.\(^{10,17,118}\)

\[(\text{Equation 6.5})\]

\[sRPE = RPE \times \text{Duration (min)}\]

**Lucia's TRIMP**

luTRIMP\(^{2,22}\) was used to calculate the TL based on HR. HR was recorded with HR monitors (Suunto, Vantaa, Finland or Pioneer, Kawasaki, Japan). luTRIMP was calculated by multiplying the time spent in 3 HR-zones according to Equation 6.6.

\[(\text{Equation 6.6})\]

\[luTRIMP = (\text{Duration (min) in zone 1} \times 1) + (\text{Duration (min) in zone 2} \times 2) + (\text{Duration (min) in zone 3} \times 3)\]

Originally the 3 HR-zones were luTRIMP is based on, are collected during laboratory exercise testing. However, no laboratory exercise testing was implemented within the team. Therefore, we determine the HR zones we used the highest HR recorded during every season and used HR zones as described by Seiler.\(^{24,85}\)
Training Stress Score
For every session, TSS was calculated according to Equation 6.1. Functional Threshold Power (FTP) was determined as 95% of the highest 20 minutes MMP obtained in every season.\textsuperscript{23,83}

Statistical analysis
Descriptive data are reported as means and SDs. An independent t-test was used to determine the significance of the differences between the mean IFs of training and race sessions. To investigate the effect of IF on the association between kJ spent and various TLs (sRPE, luTRIMP and TSS), the collected data for each individual cyclist was divided into 3 categories of IF, based on the presence of RPE, HR and PO. Each session within each individual cyclist was categorized as low, middle or high based on IF by dividing all available sessions in 3 equal groups using tertiles. When it was not possible to divide the number of sessions within an individual precisely into 3 groups with the same sample size, the “left-over” data were put in the middle group, leaving the low- and high-intensity groups with exactly the same sample size. The middle IF group was excluded from all analyses because sessions close to the median are more likely to be misclassified for IF and thus excluding the middle group enhances contrast in the analysis. Linear regression analyses were used to investigate whether training intensity (low or high) affects the association between kJ spent and the various other TLs (sRPE, luTRIMP and TSS). Generalized Estimating Equations (GEE) with an exchangeable working correlation structure were used for the regression analyses to account for the dependency of the repeated measurements within participants.\textsuperscript{127}

First, the associations between kJ Spent and each of the 3 TL measures was determined, with TL (sRPE, luTRIMP or TSS) as outcome (dependent) variable and kJ spent as predictor (independent) variable, according to the following Equation:

\[
\text{TL (sRPE, luTRIMP or TSS)} = b_0 + b_1 \times \text{kJ spent TL}
\]

with \(b_0\) representing the intercept (constant) of the regression line and \(b_1\) representing the slope (regression coefficient) for sRPE, luTRIMP or TSS. This
Chapter 6

association is affected (or modified) by IF when the regression coefficients (or slopes) of the 3 TLs (sRPE, luTRIMP or TSS) are different for the exercise sessions with a low or high IF. Statistically this would mean that the interaction (effect modification) between TL (sRPE, luTRIMP and TSS) and IF group \( b3 \) would be significant in:

\[
\text{Equation 6.8}
\]

\[
\text{TL (sRPE, luTRIMP or TSS)} = b0 + b1 \times \text{kJ spent TL} + b2 \times \text{IF group} + b3 \times \text{kJ spent TL} \times \text{IF group}
\]

Generalized estimating Equations were used to estimate the regression coefficients and their 95\% confidence intervals. All statistical analyses were performed using the Statistics Toolbox of Matlab (Release 2012b, The MathWorks, Inc., Natick, Massachusetts, United States) and IBM SPSS Statistics (version 23, IBM Corporation, Armonk, NY, USA) and a p-value < 0.05 was considered as statistically significant.

### 6.3 Results

In total, 21 professional cyclists collected PO data of 11 716 training and race sessions. HR data from 7628 sessions and RPE scores from 5460 sessions are available. A significant difference was found in the average IF between training and race sessions (0.59 [0.03] vs. 0.73 [0.03]). Based on the IF, sessions were classified as low- and high-intensity sessions. All data of the low- and high-intensity sessions are shown in Figure 6.1 and average data on these sessions are presented in Table 6.1.

**Association between kJ spent and various TLs**

The results of the regression analyses on all data with Equation 6.7 as model are presented in Table 6.2. The significant regression coefficients for sRPE, luTRIMP and TSS show that TL increases with higher levels of kJ spent for all 3 TLs (sRPE, luTRIMP and TSS).

To investigate if the association between kJ spent and sRPE, luTRIMP and TSS is affected by IF, regression coefficients for the interaction between kJ spent and IF are determined with Equation 6.8 as the regression model. Results are presented in Table 6.3. For sRPE and luTRIMP no significant regression coefficients are observed for the interaction term \( b3 \), Equation
6.8), indicating that the association between kJ spent and either sRPE or luTRIMP is not different for low and high IF sessions. This can also be seen in the upper 2 graphs of Figure 6.1. However, for TSS a significant regression coefficient (0.018) for the interaction term is found ($p < 0.001$), which means that there is a different association between kJ spent and TSS when sessions are done with a low or high IF. The predicted regression lines for low and high IF sessions are shown in Figure 6.1.

**Table 6.1:** Number of subjects (N[cyclists]) and number of observations (N[sessions]), average kJ spent, average of various TLs (sRPE, luTRIMP and TSS in AU) and average intensity factor (IF) for the low- and high-intensity group.

<table>
<thead>
<tr>
<th></th>
<th>Low IF</th>
<th>High IF</th>
<th>Low IF</th>
<th>High IF</th>
<th>Low IF</th>
<th>High IF</th>
</tr>
</thead>
<tbody>
<tr>
<td>sRPE</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>luTRIMP</td>
<td>1814</td>
<td>1814</td>
<td>2535</td>
<td>2535</td>
<td>3899</td>
<td>3899</td>
</tr>
<tr>
<td>TSS</td>
<td>1671 (1059)</td>
<td>3273 (1256)</td>
<td>1757 (1072)</td>
<td>3399 (1235)</td>
<td>1548 (1042)</td>
<td>3375 (1289)</td>
</tr>
<tr>
<td>TL (AU)</td>
<td>1776 (1221)</td>
<td>3734 (1517)</td>
<td>166 (97)</td>
<td>321 (120)</td>
<td>70 (51)</td>
<td>234 (90)</td>
</tr>
<tr>
<td>IF</td>
<td>0.51 (0.07)</td>
<td>0.74 (0.06)</td>
<td>0.53 (0.06)</td>
<td>0.76 (0.07)</td>
<td>0.51 (0.07)</td>
<td>0.77 (0.07)</td>
</tr>
</tbody>
</table>

Abbreviations: IF, intensity factor; luTRIMP, Lucia’s TRIMP; sRPE, session rating of perceived exertion; TL, training load; TSS, training stress score

**Figure 6.1:** All data of the low- and high-intensity sessions (low=black dots, high=gray dots) with the predicted regression lines for the low (black line) and high (gray line) IF sessions. Abbreviations: sRPE, session rating of perceived exertion; luTRIMP, Lucia’s TRIMP; TSS, training stress score.
Chapter 6

6.4 Discussion

This study found a significantly different average IF between training sessions and race sessions in professional cycling. Furthermore, this study found a significantly different regression coefficient for the association between kJ spent and various TLs (sRPE, luTRIMP and TSS) for sessions with a low and high exercise intensity based on IF, while there are no significant differences between the association of kJ spent and sRPE and luTRIMP for sessions with a low or high exercise intensity. This means, in practice, that there will be a significant effect on the collected TSS values for sessions done with the same amount of kJ spent, but with a low or high intensity, while there is no effect on the collected sRPE and luTRIMP. For example, as shown by van Erp et al.,\textsuperscript{103} an average race is completed with ±3700 kJ, when this race is done with a low exercise intensity, the values of Table 6.3 predict a TL of 4033, 336 and 172 AU for sRPE, luTRIMP and TSS respectively. However, when the same amount of kJ is spent in a race with high exercise intensity the presented values of Table 6.3 predict a TL of 4223 AU, 349 AU

### Table 6.2: Results of the regression analyses using General Estimation Equations, with Equation 6.7 (sRPE, luTRIMP or TSS = b0 + b1*kJ spent TL) as model, for the association between kJ spent and various TLs (sRPE, luTRIMP and TSS) for all sessions (combined low- and high-intensity sessions).

<table>
<thead>
<tr>
<th></th>
<th>Regression coefficient</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>sRPE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b0</td>
<td>-149*</td>
<td>[-239, -59]</td>
</tr>
<tr>
<td>b1</td>
<td>1.17*</td>
<td>[1.11, 0.23]</td>
</tr>
<tr>
<td>luTRIMP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b0</td>
<td>13.81*</td>
<td>[5.54, 22.07]</td>
</tr>
<tr>
<td>b1</td>
<td>0.090*</td>
<td>[0.085, 0.095]</td>
</tr>
<tr>
<td>TSS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b0</td>
<td>-23.34*</td>
<td>[-26.95, -19.72]</td>
</tr>
<tr>
<td>b1</td>
<td>0.071*</td>
<td>[0.067, 0.074]</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; luTRIMP, Lucia’s TRIMP; sRPE, session rating of perceived exertion; TL, training load; TSS, training stress score. Note: For sRPE, luTRIMP and TSS the estimated value of the regression coefficients and 95% CI are presented. * significant regression coefficients.
The Influence of Exercise Intensity on various Training Load Measures

and 255 AU for sRPE, luTRIMP and TSS. This would give a difference of 5%, 4% and 48% for sRPE, luTRIMP and TSS, respectively between low- and high-IF sessions, showing that TSS reacts different on exercise intensity in contrast to sRPE and luTRIMP.

Historically, measures of TL have a linear or exponential relationship with exercise intensity, where TSS has a quadratic relationship. A quadratic relationship seems far from obvious as no internal load unit, including RPE, relative VO\textsubscript{2} consumption, HR, blood lactate concentration, biochemical and hormonal responses, has quadratic relationship with exercise intensity. Interestingly, Sanders at al.\textsuperscript{16} demonstrated that despite the quadratic relation with exercise intensity, TSS showed a very strong dose-response relationship

### Table 6.3: Results of the regression analyses using General Estimation Equations, with Equation 6.8 (sRPE, luTRIMP or TSS = b0 + b1*KJ spent TL + b2*IF\textsuperscript{group} + b3*KJ spent TL*IF\textsuperscript{group}) as model, for the association between kJ spent and various TLs (sRPE, luTRIMP and TSS) for all sessions combined and low- and high-intensity sessions.

<table>
<thead>
<tr>
<th>Regression coefficient</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>sRPE</strong></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>-113.6* [ -183.05, -44.20]</td>
</tr>
<tr>
<td>b1</td>
<td>1.127* [1.064, 1.190]</td>
</tr>
<tr>
<td>b2</td>
<td>40.51 [-57.58, 138.59]</td>
</tr>
<tr>
<td>b3</td>
<td>0.034 [-0.029, 0.097]</td>
</tr>
<tr>
<td><strong>luTRIMP</strong></td>
<td></td>
</tr>
<tr>
<td>b0</td>
<td>13.94* [9.07, 18.81]</td>
</tr>
<tr>
<td>b1</td>
<td>0.087* [0.084, 0.091]</td>
</tr>
<tr>
<td>b2</td>
<td>12.87* [2.83, 22.91]</td>
</tr>
<tr>
<td>b3</td>
<td>0.0 [ -0.006, 0.005]</td>
</tr>
<tr>
<td><strong>TSS</strong></td>
<td></td>
</tr>
<tr>
<td>b0</td>
<td>-5.40* [-7.41, -3.39]</td>
</tr>
<tr>
<td>b1</td>
<td>0.048* [0.045, 0.050]</td>
</tr>
<tr>
<td>b2</td>
<td>15.92 [12.08, 19.85]</td>
</tr>
<tr>
<td>b3</td>
<td>0.018* [0.016, 0.020]</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; luTRIMP, Lucia’s TRIMP; sRPE, session rating of perceived exertion; TL, training load; TSS, training stress score. Note: For sRPE, luTRIMP and TSS the estimated value of the regression coefficients and 95% CI are presented. * significant regression coefficients.
between the increase in TL and PO at 2 and 4 mmol, where LuTRIMP and sRPE showed the weakest dose-response relationship. This suggests that TSS could be a more valuable measure of TL compared to sRPE and LuTRIMP in professional cycling. However, the study of Sanders et al.\textsuperscript{16} was performed in the pre-season were mainly the nature of training is low-intensity training. Because the present study shows that exercise intensity affects the kJ spent-TSS relationship, it is interesting to study the dose-response relationship of TSS during periods with more high exercise-intensity (e.g. pre-competition, competition phases). The assumption can be made that there could be a risk for athletes and coaches focusing on a polarized training program solely based on TSS. As this quadratic relation with exercise intensity could result in an underestimation of the physical impact of a long easy endurance ride compared to a short high-intensity ride, while spending the same amount of kJ.

There was no significant effect of exercise intensity on the relation between sRPE and kJ spent. This could be a reason for the weaker dose-response relationship in relation with aerobic fitness\textsuperscript{16,10} as high intensity exercise is a critical part of the training program and has a fairly rapid effect on performance.\textsuperscript{85} However, our findings are in contrast with the finding of Borresen et al.\textsuperscript{109} which suggested that the sRPE method might overestimate TL for athletes spending more time doing low-intensity exercise. Despite that professional endurance athletes spend the most of their time doing low intensity exercise, it seems that professional cyclists are more familiar with doing high intensity exercise compared to active fit individuals and therefore exercise intensity does not influence the RPE-score in professional cyclist as much as it does in active people.

Furthermore, this study shows no significant effect of exercise intensity on the relationship between kJ spent and LuTRIMP, despite that LuTRIMP is based on a summation of time spent in different intensity zones. The first reason could be that LuTRIMP is responding the same as the increase of kJ during high intensity exercise, because 1 minute in a higher HR zone means also a higher PO, which results in more kJ spent in that minute. Second, it is shown that accumulating fatigue decreases submaximal\textsuperscript{71,128} and maximum HR\textsuperscript{22} in professional cyclists. Since this study evaluates data from pre-season, pre-competitions and competitions phases (including multiple-day races) and we did not modify the HR-zones for the suppression in HR with fatigue, it
could be that luTRIMP-scores are reduced in relation to kJ spent during phases of accumulating fatigue which are most common in phases with a lot of high intensity work and races. It seems likely that fatigue will suppress luTRIMP in training or competition periods with a high load demand. Because the study of Sanders et al.\textsuperscript{16} was done in a pre-season preparatory training period, where the training in professional cycling is focused on low-intensity training, the decrease of luTRIMP with fatigue, together with the increases of TSS compared to the other load metrics at high intensity, makes it highly interesting to study the dose-response relationship between different measures of TL in the pre-competition and competition phases with a high load demand.

**Limitations**
The presented data were collected (1) for monitoring professional cyclists and (2) for the purpose of this study. This can cause some limitations on the collected data. One of the limitations could be that the FTP of each cyclist was determined by using the 20-minute MMP of the particular season. As shown by Nimmerrichter\textsuperscript{83} this value can change by 0.4 W·kg\textsuperscript{-1} during the season and therefore using the highest 20 minute MMP to calculate FTP for an entire season can cause an under- or overestimation of TSS in the different phases of the year. Furthermore, the HR zones used for luTRIMP are based on the registered maximal HR instead of HR-zones anchored on physiological thresholds obtained during incremental testing. This could have influenced the values obtained for luTRIMP. It would have been ideal to perform systematic and controlled laboratory testing on these subjects to provide this data for our analysis. It was, however, not possible to incorporate this into the collaboration with the team. Furthermore, the timing of the RPE scores was not controlled and, therefore, it could be that RPE is recorded later than the recommended 30 min after the exercise.\textsuperscript{10,17,118} However, recent data have shown that sRPE is robust relative to the timing between the exercise bout and the collection of the RPE- score.\textsuperscript{100}

**Conclusion**
This study shows that the IF is significantly different for training and race sessions. Furthermore, kJ spent is significantly associated with each of the 3 measures of TL, sRPE, luTRIMP and TSS. While the associations between kJ
spent and sRPE and luTRIMP are not influenced by IF, the association between kJ spent and TSS is different for cycling sessions with a low or a high IF. Because different measures of TL are reacting differently on exercise intensity, future research should evaluate dose-response relationships with different training load metrics, including periods with high load demands and big proportions of high intensity exercise (e.g. pre-competitions and competition phase).

**Application to practice**
A valid measure for TL is important for professional cyclists as training plans are based on TL and TL is used to monitor professional cyclists. Sports scientists and coaches working in professional cycling should be aware that TL based on PO (TSS) reacts differently for different exercise intensities, in contrast to sRPE and luTRIMP. Therefore, it could be a potential risk to focus solely on one measure of TL.
Various Workload Models are Associated with Overuse Injuries, but not with Illnesses, in Professional Female Cyclists.

Teun van Erp • Taco van der Hoorn • Marco Hoozemans • Carl Foster • Jos J. de Koning

In preparation for publication
Abstract

Objective: To determine if workload and various workload models are associated with increased likelihood of injuries and illnesses in female professional cyclists.

Methods: Data were collected from 15 elite female professional cyclists, containing a total of 22 seasons. Session Rate of Perceived Exertion (sRPE) was used to quantify internal workload. One week (acute) and four week (chronic) workload, together with three acute:chronic workload models were calculated. Two workload models are based on moving averages of the ratios, the Acute:Chronic Workload Ratio (ACWR) and the Acute:Chronic Workload Ratio Uncoupled (ACWR\textsubscript{uncoupled}), the difference between both is the chronic load. In ACWR the acute load is part of the chronic load, in ACWR\textsubscript{uncoupled} the acute and chronic load are uncoupled. The third workload model is based on Exponentially Weighted Moving Averages of the ratios (EWMA). GEE analysis was used to assess the associations between likelihood of injuries and illnesses and the various workload. To investigate if there is a long-term increased risk we analyzed the association between the workload and injuries or illnesses in the current week (week 0), the week next (week +1) and the week after next (week +2).

Results: High values of acute workload, ACWR, ACWR\textsubscript{uncoupled} and EWMA are significantly associated with a higher likelihood of injury in the current and the next week. High values of chronic workload were associated with significantly lower likelihood of injuries in the next and the week after next. No clear associations are found for the occurrence of illnesses for either workload or workload models.

Conclusions: These findings demonstrate the importance of monitoring workload and workload models in female professional cyclists to lower the risk of injuries.
7.1 Introduction

Professional cycling is a highly demanding sport where female elite cyclists typically cover between 13000 to 18000 kilometers in training and competition each year. Injuries and illnesses significantly interfere with chances of a successful season. Elite athletes that can maintain training availability at 80% have a significantly greater chance of achieving their key performance goals. Therefore in order to maximize the likelihood of success in elite athletes, attention should be paid to the prevention of both injury and illness.

Knee pain is the most common injury in elite cycling with the highest prevalence in the pre-season. In multiple sports, associations between a high workload and injuries are reported. Therefore a high workload could also be a risk factor for injuries in elite cycling. A rapid increase in workload is related to a higher likelihood of injuries, in contrast to a high chronic workload, which has been reported to be protective against injuries. Hulin et al. introduced a workload model by calculating the Acute:Chronic Workload Ratio (ACWR) and showed that high peaks in ACWR are more likely to be associated with injuries than absolute workload alone. However, recently Lolli et al. criticized the construct of ACWR because of ‘mathematical coupling’ between acute and chronic workload and suggested calculating ACWR without including the acute load as part of the chronic load (ACWR uncoup). Furthermore Menaspa raised concerns about using rolling averages to assess workload, because they do not consider the time frame in which a given stimulus occurred, nor the changing nature of fitness and fatigue over time. As a result Williams et al. proposed the use of an exponentially weighted moving average model (EWMA) and recently Murray et al. demonstrated that EWMA is more sensitive to detect increased likelihood of injury with peaks in workload compared to the ACWR model.

Comparable to injuries, similar associations are observed between workload and illnesses, although the evidence is not as clear as the relationship between workload and injury. Multiple components of the immune system exhibit changes after prolonged heavy exertion. During this so called ‘open window’ of altered immunity, which may last between 3 and 72 hours, viruses and bacteria may settle and increasing the risk of infection.
is well studied. However, the association between workload models and injuries and illnesses has not been studied in athletes participating in a high volume, but low mechanical impact sport like elite cycling. Therefore, the research question of this study was to determine the association between different workload models (acute, chronic, ACWR, ACWR\textsubscript{uncoup} and EWMA) and the likelihood of injuries and illnesses. We hypothesize that a high chronic workload protects against injuries, while high values of acute workload, ACWR, ACWR\textsubscript{uncoup} and EWMA are associated with a higher likelihood of injuries and illnesses.

### 7.2 Methods

#### Participants
In total, 15 highly trained female professional cyclists, part of a World Tour professional cycling team, participated in this study. The team finished within the top-10 of the Union Cycliste Internationale elite women’s team ranking over the course of the study period. In total we collected 22 seasons of data. The dataset of an individual cyclist ranges from one to three seasons depending on their involvement in the team and the amount of missing data. Institutional ethics approval was granted and, in agreement with the Helsinki Declaration, written informed consent was obtained from the participants.

#### Quantifying of workload
Internal workload data were collected based on rating of perceived exertion (RPE).\(^\text{18}\) All cyclists were obligated by the team to fill in their online training logbook daily, as soon as possible after their training session or race. One of the questions in the logbook was: “How hard was your day”, which was ranked on a 6-20 BORG-scale.\(^\text{18}\) The workload was calculated by multiplying the RPE of the session by the duration of the session in minutes (sRPE), according to Foster et al.\(^\text{17}\) The duration was based on the collected Power Output (PO) data of that particular day or was reported by the cyclist when PO was not recorded. Further, when no RPE-score was provided by the cyclist, but PO was recorded, we used the registered workload (kJ spent based on PO) to estimate sRPE. For every season, individual linear relationships between sRPE and kJ
spent were determined by using the Pearson-correlation and missing sRPE’s were estimated based on the regression coefficients. As shown by Van Erp et al. this relationship is almost perfect during training, racing and time-trials in professional cyclists\textsuperscript{103} and, therefore, a reliable way to limit the amount of missing data. As recommended by Hopkins, Fisher’s z-transformation was used to transform correlation coefficients (r) before averaging.\textsuperscript{79}

Data were categorized into weekly blocks running from Monday to Sunday. Acute workload corresponds to the sum of the workload for that particular week (week 0), where chronic workload was calculated as the sum of workloads for 4 weeks. Three different workload ratios were calculated: ACWR, ACWR\textsubscript{uncoup} and EWMA. ACWR is the ratio between the acute workload and chronic workload of the current and the 3 preceding weeks (week 0, -1, -2, -3). ACWR\textsubscript{uncoup} is the ratio between the acute workload and the chronic workload of the 4 preceding weeks (week -1, -2, -3, -4). The EWMA workload was calculated as described by Williams et al.\textsuperscript{135} and Murray et al.\textsuperscript{136} and presented in the following Equations:

\begin{equation}
\text{EWMA}_{\text{today}} = \text{Load}_{\text{today}} \ast \lambda_a + ((1 - \lambda_a) \ast \text{EWMA}_{\text{yesterday}})
\end{equation}

\begin{equation}
\lambda_a = \frac{2}{(N + 1)}
\end{equation}

Where is a value between 0 and 1 representing the degree of decay with N the number of days. N is chosen to be 7 days for acute and 28 days for chronic workloads. The EWMA workload ratio was calculated by dividing the EWMA acute workload by the EWMA chronic workload. The highest EWMA ratio occurring in a week was used as value for that particular week. All five workload models were corrected when an injury or illness occurred which influenced the acute workload, by using the workload 7 days prior to an injury or illness as acute workload. This was done to avoid inaccurate relationships in the current week as an injury or illness will decrease the workload in that particular week.

**Injury data collection**

The cyclists reported injuries and illnesses in an online logbook. Only nontraumatic injuries were included, injuries as a result of crashes or accidents...
were excluded. Further, illnesses not caused by an infection, like headache or menstrual pain were excluded.

**Data analysis/statistics**

To study the associations between the occurrence of injuries and illnesses, and workload we divided the workloads and workload models into four workload groups based on quartiles, representing a low (<25\(^{th}\) centile), middle-low (>25\(^{th}\) to 50\(^{th}\) centiles), middle-high (>50\(^{th}\) to 75\(^{th}\) centile) and high (>75\(^{th}\) centile) workload group.

A Generalized Estimating Equations (GEE) analysis was used (IBM SPSS Statistics version 23, IBM Corporation, Armonk, NY, USA) to determine the association between the likelihood of injuries and illnesses and the workloads and workload ratios. GEE analysis was chosen because of its ability to take into account the dependency of the repeated measurements within participants. Specifically, a binary logistic GEE was used, with a robust estimator and an exchangeable structure concerning the working correlation matrix. The odds ratio (OR), and its 95\% CI were calculated with GEE analysis and are presented together with the incidence of an injury or illness per week. For all 5 workload models the low workload group is chosen as the reference group.

To investigate if workloads and workload models have an influence over a longer period, GEE analysis was used to determine the association between workloads and workloads ratios and injuries and illness in the current week (week 0), the week next (week +1) and the week after that (week +2). The level of significance was set at (P<.05).

**7.3 Results**

In total, 22 seasons with a total of 1054 weeks of training, consisting of 5190 training days were collected. On average a season consisted of 236±30 training days. During those 22 seasons a total of 18 non traumatic injuries were reported, in which 72\% of the injuries were related to knee problems and 67\% of these cases required an adjustment to the training program. Further, 74 illnesses were reported from which 64\% required an adjustment of the training program. In total 186 training days (3.6\%) were without a reported
Training Load and the Association between Injuries and Illnesses

RPE-score. Based on the individual correlations between kJ spent and sRPE ($r=0.93\pm0.03$). We were able to estimate the sRPE-score of 136 training days based on the available kJ spent for that session, thus for 50 training days it was not possible to determine the sRPE because both kJ spent and sRPE were missing.

**Workload and workload ratios**

For GEE analysis all parameters were divided in four centiles which are presented in Table 7.1.

Table 7.1: Boundaries of the 5 workload models when divided into the four centiles.

<table>
<thead>
<tr>
<th>Workload measure</th>
<th>Low</th>
<th>Middle-low</th>
<th>Middle-high</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute load</td>
<td>&lt;7551 AU</td>
<td>7551 - 10000 AU</td>
<td>10000 - 12877 AU</td>
<td>&gt;12877 AU</td>
</tr>
<tr>
<td>Chronic load</td>
<td>&lt;8561 AU</td>
<td>8561 - 10372 AU</td>
<td>10372 - 11912 AU</td>
<td>&gt;11912 AU</td>
</tr>
<tr>
<td>ACWR</td>
<td>&lt;0.77</td>
<td>0.77 – 0.99</td>
<td>0.99 – 1.28</td>
<td>&gt;1.28</td>
</tr>
<tr>
<td>ACWR uncoup</td>
<td>&lt;0.74</td>
<td>0.74 – 0.99</td>
<td>0.99 – 1.4</td>
<td>&gt;1.42</td>
</tr>
<tr>
<td>EWMA</td>
<td>&lt;1.10</td>
<td>1.10 – 1.23</td>
<td>1.23 – 1.38</td>
<td>&gt;1.38</td>
</tr>
</tbody>
</table>

Abbreviations: ACWR, acute chronic workload ratio, ACWR uncoup, acute chronic workload ratio uncoupled ratio, EWMA, exponentially weighted moving average ratio.

GEE analysis showed significant associations between a high acute workload and a higher likelihood of injuries compared to a low acute workload in the current and the next week. Thus, cyclists with a high acute load have a 4.5 (95%CI=1.2 to 16.8) and 3.7 (95%CI=1.1 to 12.0) times higher change of an injury in the current and next week, respectively, compared to a low acute load. No significant associations were found between a high acute workload compared to a low acute workload and injuries in the week after next (Table 7.2).

Further the GEE analysis showed significant associations between a lower likelihood of injuries with a high chronic workload compared to a low acute workload in the next week. Thus, cyclists with a high chronic load have a 0.25 (95%CI 0.1 to 0.8) times higher change of an injury in the next week compared to cyclists with a low chronic load. Thus, cyclists with a low chronic load have a 4.0 times higher change of an injury compared to cyclists with a high chronic load. In addition, a lower likelihood of injuries was found in the week after next with a low chronic workload compared to a middle-high chronic workload (OR=0.1, 95%CI=0.05 to 0.36) (Table 7.2). In general, a
middle-high to high chronic load seems to be protective for injury.

GEE analysis showed a significant association between an increased likelihood of injuries and a high ACWR\textsubscript{uncoup} compared to a low ACWR\textsubscript{uncoup} in the current week (OR=9.6, 95%CI=1.2 to 85.3), while a high ACWR (p=0.07) and a high EWMA (p=0.07) had the tendency to have an increase likelihood of injuries. In addition, for all the three workload models (ACWR, ACWR\textsubscript{uncoup} and EWMA) a significant increased likelihood of injuries was found between a high workload compared to a low workload in the next week (OR=10.1, 95%CI=1.8 to 58.0, OR=10.3, 95%CI=1.5 to 69.4, and OR=16.1, 95%CI=2.5 to 33.3 for ACWR, ACWR\textsubscript{uncoup} and EWMA, respectively). No significant associations were found for the workload models in the week after next and the likelihood of injuries. However, in the EWMA low workload group no injuries were reported and therefore statistical testing was not possible (Table 7.2).

For illnesses only a significant association (OR=1.9, 95%CI=1.1 to 3.2) was found for a higher likelihood of illnesses in the next week in the middle-low ACWR\textsubscript{uncoup} group compared to the low ACWR\textsubscript{uncoup} group, while no other significant associations were found (Table 7.3).

### 7.4 Discussion

To our knowledge this is the first study to investigated acute workload, chronic workload and three different workload models and their associations with injuries and illnesses in female professional cyclists.

**Acute and chronic workload and the likelihood of injuries**

Our results show a fourfold increased risk of injuries (mostly knee injuries) with a high acute workload in elite female cycling in the current week and next week. Further, this study shows that a high chronic workload is protective against injuries. With ORs of 0.25 and 0.1 for the middle-high and high chronic workload groups, indicating that the low chronic workload groups have a 4 (1/0.25) to 10 (1/0.1) times higher risk of injuries in the weeks after exposure compared to the high chronic workload groups. Both findings are in line with the literature.\textsuperscript{14,131} Additional statistical testing revealed a significant lower likelihood of injuries in the middle-high workload group compared to
the middle-low (OR=0.09, 95%CI 0.01 to 0.60) and the low workload group (OR=0.13, 95%CI 0.05 to 0.36). However, no significant associations were found between a lower likelihood of injuries in the high chronic workload group compared with the low chronic workload group in the week after next, which is probably caused by the low number of injuries.

**Differences in workload models**

Currently it is debated which workload models are best to use for injury and illness prevention.\textsuperscript{133,134,138} The ACWR workload model is evidence based and strongly supported by the available literature.\textsuperscript{14} Nevertheless, this model has been criticized for the fact that the numerator and denominator are mathematically coupled and it is suggested that rolling averages do not consider when a given stimulus happened within the set time frame.\textsuperscript{133,134} To investigate the claim that alternative workload models are better for injury and illness prevention, we analysed three different workload models. Our data showed that a high ACWR\textsubscript{uncoup} has the highest association with the likelihood of injuries in the current week, up to a tenfold, compared to a high ACWR and a high EWMA, which have a fivefold higher likelihood (not significant). In the week prior to injury all three models found significant associations with an increased likelihood of injuries next week compared with a low workload. However, ACWR and ACWR\textsubscript{uncoup} had higher odds ratios compared to the EWMA workload, which is in contrast with the results of Murray et al.,\textsuperscript{136} who found that EWMA is more sensitive to detect injury risk compared to ACWR in elite Australian football athletes. Although there is some variation in odds-ratios for the different models, the count of injuries did not differ by more then 2 injuries between the different workload models. Therefore, it is difficult to make conclusions. One of the reasons for these small differences could be caused by the limited number of injuries (18) found in this study, despite analysing over 1054 weeks of training data. As suggested by Twisk\textsuperscript{139} it is recommended that for normal distributions a minimal of 5 injuries should be present in a workload group. Therefore, we could only distinguish between four different workload groups and not between more extreme workload ratios (e.g. <0.49 or >2.0) in contrast to the a previous study.\textsuperscript{136} This could be a reason for not finding more striking differences between the different workload models for predicting the likelihood of injuries.
Table 7.2: Incidence and associations between injuries in the current week, the next week and the week after next between the four different workload categories (I: <25<sup>th</sup> (low), II: >25<sup>th</sup> to 50<sup>th</sup> (middle-low), III: >50<sup>th</sup> to 75<sup>th</sup> (middle-high), VI: >75<sup>th</sup> (high)).

<table>
<thead>
<tr>
<th>Workload-model</th>
<th>Injuries</th>
<th>Current week (week 0)</th>
<th>Next week (week +1)</th>
<th>Week after next (week +2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Odds Ratio (95%CI)</td>
<td>P value</td>
<td>Incidence</td>
</tr>
<tr>
<td>Acute load</td>
<td>I</td>
<td>2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>3</td>
<td>1.5 (0.3 to 8.1)</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>4</td>
<td>2.0 (0.4 to 9.9)</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>VI</td>
<td>9</td>
<td>4.5 (1.2 to 16.8)</td>
<td>0.03*</td>
</tr>
<tr>
<td>Chronic load</td>
<td>I</td>
<td>5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>3</td>
<td>0.6 (0.2 to 1.9)</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>3</td>
<td>0.5 (0.2 to 1.6)</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>VI</td>
<td>7</td>
<td>1.3 (0.7 to 2.4)</td>
<td>0.48</td>
</tr>
<tr>
<td>ACWR</td>
<td>I</td>
<td>2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>0</td>
<td>#</td>
<td>#</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>7</td>
<td>1.3 (0.7 to 2.6)</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>VI</td>
<td>9</td>
<td>4.5 (0.89 to 21.1)</td>
<td>0.08</td>
</tr>
</tbody>
</table>

Table continued on the next page
### Training Load and the Association between Injuries and Illnesses

**Chapter 7**

<table>
<thead>
<tr>
<th>Workload-model</th>
<th>Current week (week 0)</th>
<th>Injuries</th>
<th>Next week (week +1)</th>
<th>Week after next (week +2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Odds Ratio (95%CI)</td>
<td>P value</td>
<td>Incidence</td>
</tr>
<tr>
<td>I</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>0.4%</td>
</tr>
<tr>
<td>II</td>
<td>1</td>
<td>1.0 (0.1 to 15.1)</td>
<td>0.97</td>
<td>0.4%</td>
</tr>
<tr>
<td>III</td>
<td>6</td>
<td>5.9 (0.7 to 51.6)</td>
<td>0.11</td>
<td>2.3%</td>
</tr>
<tr>
<td>VI</td>
<td>10</td>
<td>9.6 (1.2 to 85.3)</td>
<td>0.04*</td>
<td>3.8%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Workload-model</th>
<th>Current week (week 0)</th>
<th>Injuries</th>
<th>Next week (week +1)</th>
<th>Week after next (week +2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Odds Ratio (95%CI)</td>
<td>P value</td>
<td>Incidence</td>
</tr>
<tr>
<td>I</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>0.8%</td>
</tr>
<tr>
<td>II</td>
<td>2</td>
<td>1.0 (1.4 to 7.7)</td>
<td>0.98</td>
<td>0.8%</td>
</tr>
<tr>
<td>III</td>
<td>5</td>
<td>2.5 (0.5 to 13.1)</td>
<td>0.27</td>
<td>1.9%</td>
</tr>
<tr>
<td>VI</td>
<td>9</td>
<td>4.6 (0.9 to 24.1)</td>
<td>0.07</td>
<td>3.4%</td>
</tr>
</tbody>
</table>

**Abbreviations:** ACWR, acute chronic workload ratio, ACWR<sub>uncoup</sub>, acute chronic workload ratio uncoupled ratio, EWMA, Exponentially weighted moving average ratio, N (number of injuries and illnesses), 95% CI, p level and Incidence. * OR are significant different to the reference group. - Reference group. # no injuries reported therefore not suitable for statistical testing.
Table 7.3 Incidence and associations between illnesses in the current week, the next week and the week after next between the four different workload categories (I: <25th (low), II: >25th to 50th (middle-low), III: >50th to 75th (middle-high), IV: >75th (high)).

<table>
<thead>
<tr>
<th>Workload-model</th>
<th>Illnesses</th>
<th>Current week (week 0)</th>
<th>Next week (week +1)</th>
<th>Week after next (week +2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Odds Ratio (95% CI)</td>
<td>P value</td>
<td>Incidence</td>
</tr>
<tr>
<td>Acute load</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>16</td>
<td>-</td>
<td>-</td>
<td>6.4%</td>
</tr>
<tr>
<td>II</td>
<td>23</td>
<td>1.4 (0.8 to 2.2)</td>
<td>0.20</td>
<td>8.6%</td>
</tr>
<tr>
<td>III</td>
<td>18</td>
<td>1.0 (0.5 to 2.0)</td>
<td>1.0</td>
<td>6.6%</td>
</tr>
<tr>
<td>VI</td>
<td>16</td>
<td>0.9 (0.6 to 1.5)</td>
<td>0.7</td>
<td>5.9%</td>
</tr>
<tr>
<td>Chronic load</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>17</td>
<td>-</td>
<td>-</td>
<td>6.8%</td>
</tr>
<tr>
<td>II</td>
<td>22</td>
<td>1.2 (0.7 to 2.0)</td>
<td>0.55</td>
<td>8.1%</td>
</tr>
<tr>
<td>III</td>
<td>14</td>
<td>0.8 (0.5 to 1.3)</td>
<td>0.33</td>
<td>5.5%</td>
</tr>
<tr>
<td>VI</td>
<td>20</td>
<td>1.0 (0.6 to 1.7)</td>
<td>0.97</td>
<td>7.0%</td>
</tr>
<tr>
<td>ACWR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>14</td>
<td>-</td>
<td>-</td>
<td>5.5%</td>
</tr>
<tr>
<td>II</td>
<td>24</td>
<td>1.7 (0.8 to 3.5)</td>
<td>0.16</td>
<td>9.0%</td>
</tr>
<tr>
<td>III</td>
<td>21</td>
<td>1.4 (0.8 to 2.6)</td>
<td>0.27</td>
<td>7.5%</td>
</tr>
<tr>
<td>VI</td>
<td>14</td>
<td>1.0 (0.5 to 1.9)</td>
<td>0.96</td>
<td>5.4%</td>
</tr>
</tbody>
</table>

Table continued on the next page
### Table continued on the previous page

<table>
<thead>
<tr>
<th>Workload-model</th>
<th>Current week (week 0)</th>
<th>Illnesses</th>
<th>Next week (week +1)</th>
<th>Week after next (week +2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Odds Ratio (95%CI)</td>
<td>P value</td>
<td>Incidence</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ACWR</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>16</td>
<td>-</td>
<td>-</td>
<td>6.4%</td>
</tr>
<tr>
<td>II</td>
<td>21</td>
<td>1.2 (0.7 to 2.3)</td>
<td>0.51</td>
<td>7.7%</td>
</tr>
<tr>
<td>III</td>
<td>22</td>
<td>1.3 (0.7 to 2.5)</td>
<td>0.40</td>
<td>8.2%</td>
</tr>
<tr>
<td>VI</td>
<td>13</td>
<td>0.8 (0.4 to 1.4)</td>
<td>0.76</td>
<td>5.1%</td>
</tr>
<tr>
<td><strong>EWMA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>17</td>
<td>-</td>
<td>-</td>
<td>6.6%</td>
</tr>
<tr>
<td>II</td>
<td>21</td>
<td>1.2 (0.5 to 2.7)</td>
<td>0.63</td>
<td>8.0%</td>
</tr>
<tr>
<td>III</td>
<td>20</td>
<td>1.1 (0.6 to 2.3)</td>
<td>0.73</td>
<td>7.5%</td>
</tr>
<tr>
<td>VI</td>
<td>15</td>
<td>0.8 (0.5 to 1.4)</td>
<td>0.50</td>
<td>5.6%</td>
</tr>
</tbody>
</table>

**Abbreviations:** ACWR, acute chronic workload ratio, ACWR\textsubscript{uncoup}, acute chronic workload ratio uncoupled ratio, EWMA, Exponentially weighted moving average ratio, N (number of injuries and illnesses), 95% CI, p level and Incidence. * OR are significant different to the reference group. - Reference group.
Ilnesses
This study found an association between a higher likelihood of illnesses in the week after exposure in the middle-low workload group. This is an unexpected result, because our hypothesis was that high acute workload causes an “open window” where the immune system is weakened. A reasonable explanation could be the incubation time of some viruses which can be 5.6 (CI 4.8 to 6.3) days.\(^{140}\) Therefore it could be that the cyclists already experience some minor symptoms, which lowers the workload in the week prior to getting sick. This is in contrast with the suggestions of Foster et al.\(^{66}\) that illnesses typically follow a spike in training load by several days.

Further, we found no clear evidence that peaks in workload or workload models increase the likelihood of illnesses in elite female cyclists. This is in contrast with the moderate relationship found by Drew & Finch\(^{14}\) and the hypothesis of the “open window” after exercise.\(^{137}\) However, recently Campbell et al.\(^{141}\) discussed the “open window” and stated that there is limited evidence and instead it is more likely that exercise improves the immune system.

Limitations
This was a retrospective study were 15 elite female cyclists with a combined 22 seasons of data were analysed. These cyclists were selected based on their consistent training logbooks. This resulted in a data-set with a limited amount of missing data, from 186 training days (3,6%) no RPE-scores were reported. However, based on the collected PO we were able to estimate the sRPE for 136 training days, thus from only 50 training days there was a missing sRPE. Although we analysed 1054 weeks of training data, we only found 18 injuries, which is limited compared to previous studies. Because of the limited amount of injuries, we could not refine the workloads into more detailed groups. Further, in this study the workloads are only based on load during cycling. Many professional cyclists are doing strength training and core stability. However, the cyclists participating in this study were not instructed to record their strength training in detail and therefore we could not include this workload into the analysis. Furthermore, the analysed injuries and illnesses were all self-reported in an online-logbook which could result in missing values.
Conclusion
This is the first study to investigate workload (models) and their associations with injuries and illnesses in elite female cycling. The findings demonstrate the importance of monitoring workload (models) in female professional cyclists to lower the likelihood of injuries, although no relations are found between workload (models) and the likelihood of illnesses.
Chapter 8

General Discussion
The aim of the present thesis was to reach a better understanding of the theoretical framework proposed by Impellizzeri for professional male and female cycling (Figure 8.1) which is of importance for coaches and sport scientists working in professional cycling. To make a detailed training program, the first step is to understand the determinants of competition and performance in professional cycling races. Therefore, we investigated the performances necessary to compete for the victory in a GT (Chapter 2) and the load and intensity demands in races and GTs (Chapters 2 and 3). In professional cycling, one of the determinants is the ability to sustain high loads with high intensity for multiple days, without any rest days. Knowledge about the load and intensity demands in races should give coaches the required knowledge to define training goals. Therefore, it is of interest to compare load and intensity demands in races with the load and intensity in training and discuss the possibilities for improvement.

The second step for coaches is to develop a training program based on the training goals of an individual cyclists. Coaches can manipulate the training program based on the FITT principles to achieve different training goals within a training session or training period. A training session or a training period contains a certain amount of TL. TL is one of the most important measures in elite sports. The relationships between external and internal TL is highly important (Figure 8.1) because the internal TL ultimately determines the functional outcome of training (i.e. positive, or negative, adaptation). To provide more knowledge about the different load measures used in cycling we described the relationship between various load measurements (i.e. LuTRIMP, sRPE, TSS and kJ spent) in training, races and TTs (Chapter 5). In addition, we investigated whether the differences between TSS and the other load measurements are caused by the quadratic relationship between TSS and exercise intensity (Chapter 6). Measuring TL with sufficient accuracy and precision is highly important for coaches and sport scientist working in professional cycling. Therefore, various investigated TL measure will be discussed in the following part of this thesis. The advantages and disadvantages of TL measure will be discussed to provide coaches and sport scientist with valuable information.
8.1 The training process in professional cycling

The load and intensity demands of professional cyclists in training and races are described in Chapters 3 and 4 and the difference between men and women are highlighted. It was hypothesized that differences in race formats could lead to differences in load and intensity between male and female cyclists. In agreement with our hypothesis, it seems that the load and intensity demands differ between men’s and women’s races (Chapter 3). Furthermore, it was suggested that different demands in races, could lead to different training programs (Chapter 4). In Chapter 4 it was concluded that a women’s training program differs from a men’s training program. It seems that women compensate their lower training volume with a higher training intensity, similar to the demands in races. The differences between the training and race demands of professional male and female cyclists are extensively discussed in Chapters 3 and 4.

Figure 8.1: Theoretical framework between training goals, external and internal training load and the training outcome. Adapted from Impellizzeri.
Knowledge about the determinants necessary to compete in races could lead to different insights for training goals. As stated in the introduction of this thesis, one of the determinants for a successful professional cyclist is to have the capacity to sustain high loads and high intensities in races. Consequently, training goals must match the demands of the races. Therefore, it is interesting to investigate if the training goals of professional cyclists match the load and intensity demands of the race (Figure 8.1). Comparing the demands of racing and training shows that the effect sizes of the differences in volume and load demands are moderate to very large when comparing in both men’s and women’s races to training (Tables 8.1, 8.2, 8.3 and 8.4). This is no surprise because part of the training sessions are recovery rides or build up sessions after a rest period. Therefore, these rides are performed with a low volume and load. Further, it was hypothesized that competing for the GC in a GT could lead up to higher demands compared to an average race. However, the data presented in this thesis showed only trivial to small effect sizes for differences between the load and intensity demands in a GT and an average race (Tables 8.1, and 8.2). Surprisingly, the effect sizes for differences of some volume and load measures (i.e. distance, duration, kJ spent and TSS) were small in a GC contender in a GT when comparing to an average race (Tables 8.1 and 8.3). A reason could be the higher physical capacity of the GC contender. The average FTP of the male cyclists from which the data are presented in Chapters 3 and 4 was 389 W or 5.3 W·kg⁻¹, while the FTP of the GC contender was 411 W or 5.9 W·kg⁻¹. The same PO in a race will result in lower load measures, (i.e. mean PO, Intensity Factor, kJ spent and TSS) for the GC contender compared to the average professional cyclist. This is clearly visible in Table 8.1 where it is shown that the average GT stage is performed at a higher average PO but with a lower TSS. In addition, the lower load measures in GT could be the result of long-term fatigue as racing over multiple days without a rest-day might influence the PO and thus the load.

When load is expressed relative to distance (i.e. eTRIMP·km⁻¹, sRPE·km⁻¹ and kJ·km⁻¹), trivial to small effect sizes were observed for difference found between the demands in training and racing (Tables 8.3 and 8.4). Further, trivial effect sizes were observed for difference found between the demands in races compared to GTs. Interestingly, except for TSS·km⁻¹, for the other load
**Table 8.1**: Volume and intensity demands in men's professional training, race and grand tours.

<table>
<thead>
<tr>
<th></th>
<th>Training Mean (SD)</th>
<th>Races Mean (SD)</th>
<th>GTs Mean (SD)</th>
<th>Training versus Races Effect sizes</th>
<th>Training versus GTs Effect sizes</th>
<th>Races versus GTs Effect sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Distance (km)</strong></td>
<td>91.9 (48.1)</td>
<td>183.0 (32.0)</td>
<td>173.0 (55.9)</td>
<td>2.27 – very large</td>
<td>1.56 – large</td>
<td>0.23 – small</td>
</tr>
<tr>
<td><strong>Duration (min)</strong></td>
<td>181.5 (95.6)</td>
<td>285.0 (56.0)</td>
<td>263.8 (90.6)</td>
<td>1.37 – large</td>
<td>0.88 – moderate</td>
<td>0.29 – small</td>
</tr>
<tr>
<td><strong>Mean PO (W)</strong></td>
<td>191.1 (32.2)</td>
<td>216.0 (34.0)</td>
<td>224.1 (42.1)</td>
<td>0.75 – moderate</td>
<td>0.89 – moderate</td>
<td>0.21 – small</td>
</tr>
<tr>
<td><strong>Mean PO (W·kg⁻¹)</strong></td>
<td>2.64 (0.45)</td>
<td>3.00 (0.50)</td>
<td>3.20 (0.60)</td>
<td>0.75 – moderate</td>
<td>1.06 – moderate</td>
<td>0.36 – small</td>
</tr>
<tr>
<td><strong>Intensity Factor™</strong></td>
<td>0.59 (0.09)</td>
<td>0.73 (0.08)</td>
<td>0.72 (0.09)</td>
<td>1.56 – large</td>
<td>1.36 – large</td>
<td>0.12 – small</td>
</tr>
<tr>
<td><strong>Mean HR (beats·min⁻¹)</strong></td>
<td>124.9 (12.4)</td>
<td>133.0 (12.0)</td>
<td>-</td>
<td>0.66 – moderate</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Mean HR (%HRmax)</strong></td>
<td>65.5 (6.3)</td>
<td>69.0 (6.0)</td>
<td>-</td>
<td>0.57 – small</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>HRmax (beats·min⁻¹)</strong></td>
<td>164.5 (18.4)</td>
<td>180.0 (12.0)</td>
<td>-</td>
<td>1.02 – moderate</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Mean RPE (6-20 scale)</strong></td>
<td>12.2 (2.30)</td>
<td>15.4 (2.10)</td>
<td>-</td>
<td>1.47 – large</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Abbreviations: GT, Grant Tour; PO, Power Output; HR, Heart Rate; RPE, rating of perceived exertion. Note: Qualitative interpretation of Cohen’s d was based on the guidelines provided by Hopkins et al.⁹⁵
measures expressed relative to distance (i.e. eTRIMP∙km$^{-1}$, sRPE∙km$^{-1}$ and kJ∙km$^{-1}$) trivial to small effect sizes were found when comparing training and races in both male and female professional cyclists. That TSS∙km$^{-1}$ contrasts with the other load measurements is not a surprise. In Chapter 6 it was shown that TSS is non-linearly influenced by exercise intensity and thus TSS behaves differently on exercise intensity in comparison with other load measures. This non-linear behavior makes it “easier” to collect TSS points in racing compared to training because the intensity in races is higher compared to training and thus TSS∙km$^{-1}$ is higher in races compared to training. Although effect sizes were trivial to small, it is interesting that the other load measures relative to distance (eTRIMP∙km$^{-1}$, sRPE∙km$^{-1}$ and kJ∙km$^{-1}$) were higher in training compared to races.

The presented data in this thesis are collected within one professional cycling team and within a team where a polarized training approached is implemented, as proposed by Seiler et al.$^{24,25,90}$ A polarized training approached is characterized by a intensity distribution that consist of ~80% of training spent in zone 1 and the remaining ~20% at zone 3, with little or none in zone 2. In multiple endurance sports (i.e. middle-distance running,$^{29,32}$ long

| Table 8.2: Volume and intensity demands in Women’s professional training and races. |
|---------------------------------|-----------------|-----------------|-----------------|
|                                 | Training        | Races           | Effect sizes between: |
|                                 | Mean (SD)       | Mean (SD)       | Training and races  |
| **Distance (km)**               | 62.6 (38.1)     | 116.0 (17.0)    | 1.9 – Large       |
| **Duration (min)**              | 144.9 (72.4)    | 194.0 (30.0)    | 1.0 – Moderate    |
| **Mean PO (W)**                 | 137.8 (22.0)    | 167.0 (21.0)    | 1.4 – Large       |
| **Mean PO (W∙kg$^{-1}$)**       | 2.30 (0.37)     | 2.80 (0.40)     | 1.3 – Large       |
| **Intensity Factor™**           | 0.66 (0.12)     | 0.83 (0.07)     | 1.8 – Large       |
| **Mean HR (beats ∙ min$^{-1}$)**| 136.2 (15.6)    | 152.0 (13.0)    | 1.1 – Moderate    |
| **Mean HR (%HRmax)**            | 69.8 (6.67)     | 79.0 (10.0)     | 1.1 – Moderate    |
| **HRmax (beats ∙ min$^{-1}$)**  | 170.4 (18.9)    | 185.0 (10.0)    | 1.0 – Moderate    |
| **Mean RPE**                    | 12.1 (2.23)     | 15.4 (1.50)     | 1.8 – Moderate    |

Abbreviations: PO, Power Output; HR, Heart Rate; RPE, rating of perceived exertion. Note: Qualitative interpretation of Cohen’s $d$ was based on the guidelines provided by Hopkins et al.$^{95}$
distance running,\textsuperscript{29,30} cross-country skiing,\textsuperscript{31,32} biathlon\textsuperscript{31} and cycling\textsuperscript{32}) there is evidence that a polarized training approach results in larger improvements in elite endurance athletes compared to other training approaches. Although the actual intensity in those sports are around threshold, it seems that a polarized training approach is more effective than a threshold training approach (in zone 2).\textsuperscript{24,29} The higher load measures expressed relative to distance in training could indicate that training within this team should be done at a lower intensity to accomplishing an even more polarized approach. In addition, this is strengthened by the observed differences in (%) time spent in the different intensity zones between training and races (Figure 8.1).

Table 8.3: Absolute load and load values expressed relative to distance in male professional cyclists during training, racing and grand tours.

<table>
<thead>
<tr>
<th></th>
<th>Training</th>
<th>Races</th>
<th>GTs</th>
<th>Training and races</th>
<th>Training and GTs</th>
<th>Races and GTs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Cohen’s d</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total work (kJ)</strong></td>
<td>2151 (1234)</td>
<td>3734 (918)</td>
<td>3653 (1240)</td>
<td>1.47 – large</td>
<td>1.21 – large</td>
<td>0.08 – trivial</td>
</tr>
<tr>
<td><strong>TSS (AU)</strong></td>
<td>115.5 (75.2)</td>
<td>255.0 (68.0)</td>
<td>229.2 (84.0)</td>
<td>1.95 – large</td>
<td>1.43 – large</td>
<td>0.34 – small</td>
</tr>
<tr>
<td><strong>eTRIMP (AU)</strong></td>
<td>453 (263)</td>
<td>739 (203)</td>
<td>–</td>
<td>1.23 – large</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td><strong>sRPE (AU)</strong></td>
<td>2495 (1437)</td>
<td>4370 (1144)</td>
<td>–</td>
<td>1.45 – large</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td><strong>kJ·km\textsuperscript{-1} (kJ)</strong></td>
<td>23.6 (6.88)</td>
<td>20.4 (3.90)</td>
<td>20.9 (10.0)</td>
<td>0.59 – small</td>
<td>0.32 – small</td>
<td>0.07 – trivial</td>
</tr>
<tr>
<td><strong>TSS·km\textsuperscript{-1} (AU)</strong></td>
<td>1.23 (0.47)</td>
<td>1.40 (0.31)</td>
<td>1.33 (0.68)</td>
<td>0.43 – small</td>
<td>0.17 – trivial</td>
<td>0.14 – trivial</td>
</tr>
<tr>
<td><strong>eTRIMP·km\textsuperscript{-1} (AU)</strong></td>
<td>4.58 (1.75)</td>
<td>4.08 (0.94)</td>
<td>–</td>
<td>0.37 – small</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td><strong>sRPE·km\textsuperscript{-1} (kJ)</strong></td>
<td>25.0 (5.52)</td>
<td>24.0 (4.30)</td>
<td>–</td>
<td>0.21 – small</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Abbreviations: GT, Grand Tour; TSS, Training Stress Score; eTRIMP, Edwards’ training impulse; sRPE, session Rating of Perceived Exertion. Note: Qualitative interpretation of Cohen’s d was based on the guidelines provided by Hopkins et al.\textsuperscript{95}
As shown in Figure 8.2, the PO intensity zones in races, especially in women’s races, are distributed in a polarized manner. It could be argued that in a sport with polarized demands in races, a polarized training approach would be preferred. In both the men’s and women’s training programs moderate effect sizes are found when comparing the amount of time spent in intensity zone 2 during training and races with higher values observed during the races. Further, a moderate to large effect sizes are found when the amount of time spent in the high intensity zones (i.e. PO zones 4 and 5) in training session is compared to the time spent in the high intensity zones during races with a again higher values observed during the races. The presented data suggest that spending more time in zone 1 to the expense of time in zone 2 and an increased time spent in the high intensity zones (i.e. PO zones 4 and 5) could lead to more effective training. This will result in a more polarized training approach and a better match with the demands of races. Although knowledge about the load and intensity demands in races is important to determine training goals, the success of the training program depends on the stimulus imposed on the individual cyclist and the essentially positive training adaptations. Further it could be that the less percent time spent in the high intensity zone (i.e. zones 4 and 5), especially in the male cyclists, is caused by implementing races as part of the training program, which is common practice in professional cycling.

Figure 8.2: Intensity distribution for men and women for %time spent in different PO zones in training versus racing (and Grand Tours) in professional cyclists. PO indicates power output. Note: Qualitative interpretation of Cohen’s d was based on the guidelines provided by Hopkins et al. *moderate (d≥0.60) effect sizes between training vs races. #moderate (d≥0.60) effect sizes between training vs grand tours. (power output data in races not presented in Chapter 3)
A disbalance within the training program between load and recovery could result in negative adaptations. A disbalance within the training program could be a rapid increase in TL. Therefore, in Chapter 7 the association between injuries and illnesses and a rapid increase in TL were investigated. It was shown that an increase of TL by more than 30% (workload ratios of >1.3) compared to the average TL for the 4 preceding weeks could result in an increased risk on nontraumatic injuries. Coaches should be aware of this increased risk when developing a training program for their athlete. One of the determinants of a successful season is to maintain the training availability above 80%. However, in Chapter 7 we found a limited amount of injuries (17) within 1054 training weeks in professional female cyclists. Furthermore, even when the TL increased with 30% compared to 4 preceding weeks a proximally 10 injuries occur within 250 training weeks. This means that increasing the TL by 30% or more, compared to the 4 preceding weeks, is without consequences for 240 out of 250 weeks. Although it is not recommended, it could be a strategy for coaches to deliberately increase the TL by more then 30% to reach certain training goals. In Chapter 7 we found a 10-fold increase in the likelihood of injuries when TL increased with more then 30%. However, in practice it means that an injury will occur in 4% of the training weeks even though the

**Table 8.4:** Absolute load and load values expressed relative to distance in female professional cyclists during training and races.

<table>
<thead>
<tr>
<th></th>
<th>Training</th>
<th>Races</th>
<th>Effect sizes between Training and races:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total work (kJ)</strong></td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Cohen’s d</td>
</tr>
<tr>
<td></td>
<td>1223 (645)</td>
<td>1958 (384)</td>
<td>1.4 – large</td>
</tr>
<tr>
<td><strong>TSS (AU)</strong></td>
<td>111 (67.7)</td>
<td>224.0 (49.0)</td>
<td>1.9 – large</td>
</tr>
<tr>
<td><strong>eTRIMP (AU)</strong></td>
<td>425 (221)</td>
<td>700.0 (141.0)</td>
<td>1.5 – large</td>
</tr>
<tr>
<td><strong>sRPE (AU)</strong></td>
<td>1851 (1070)</td>
<td>2982 (585)</td>
<td>1.4 – large</td>
</tr>
<tr>
<td><strong>kJ·km⁻¹ (kJ)</strong></td>
<td>17.7 (2.42)</td>
<td>16.8 (2.50)</td>
<td>0.38 – small</td>
</tr>
<tr>
<td><strong>TSS·km⁻¹ (AU)</strong></td>
<td>1.56 (0.49)</td>
<td>1.92 (0.35)</td>
<td>0.86 – moderate</td>
</tr>
<tr>
<td><strong>eTRIMP·km⁻¹ (AU)</strong></td>
<td>5.72 (1.23)</td>
<td>6.02 (0.84)</td>
<td>0.29 – small</td>
</tr>
<tr>
<td><strong>sRPE·km⁻¹ (kJ)</strong></td>
<td>25.9 (4.95)</td>
<td>25.6 (4.00)</td>
<td>0.07 – trivial</td>
</tr>
</tbody>
</table>

Abbreviations: TSS, Training Stress Score; eTRIMP, Edwards’ training impulse; sRPE, session rating of perceived exertion. Note: Qualitative interpretation of Cohen’s d was based on the guidelines provided by Hopkins et al.
Chapter 8

TL is increased by more than 30% compared to the preceding weeks. Thus, it could be worth the risk in some circumstances. The results from Chapter 7 give coaches the knowledge to make an informed choice if this is a risk they are willing to take.

8.2 External and Internal Training load in Professional Cycling

sRPE

sRPE is an easy to use, noninvasive and affordable TL measure, which can be used for monitoring athletes.\textsuperscript{10,17} sRPE is a highly reliable tool to assess TL, provided that coaches and athletes are motivated to assess daily TL, because no equipment is necessary and thus there is no risk of malfunctions. In Chapter 7 the cyclists were instructed to collect power-data and RPE-scores. Although the group of cyclists was highly motivated there was still some missing data. From all training sessions, 16.3% of the PO data were missing due to malfunctions or riding on bikes which were not equipped with a power meter (i.e. mountain bike or TT bike), while only 3.6% of the RPE-scores were missing. In Chapter 2, in total 8 out of 64 power files were missing caused by malfunctions of the power meter or crashes. Although easy to use, collection of sRPE requires discipline from the athletes and coaches in order to collect daily TL. When the discipline is lacking, a power or HR monitor will be easier to collect daily TL, mostly because the data are stored without any action of the athletes. This could be the reason that the datasets used in Chapters 3, 4, 5 and 6 contain a higher amount of sessions with PO data compared the number of RPE values. Thus, the strength of the use of sRPE depends on the coaches and athletes. When motivated there is no reason for missing data. Because of the exceptionally strong data set collected in Chapter 7 we were able to investigate the associations between sRPE and the occurrence of injuries and illnesses and it was shown that sRPE is a valuable tool to monitor professional cyclists.

Chapter 5 showed that sRPE is highly associated with other measures of TL (i.e. luTRIMP, kJ spent and TSS) in training, races and TTs. However, the association was slightly weaker during races, which is probably caused by factors that are less controllable in races (i.e. hydration\textsuperscript{115} and environmental
Another reason could be that fatigue influences RPE-scores, and a stage-race consists of multiple stages without a rest day and thus accumulated fatigue could influence RPE-scores. As suggested by Sanders et al., the relationship between RPE (sRPE) and other measurements of TL could be used to assess fatigue in professional cyclists. The sRPE:TSS ratio was moderately higher during a GT compared to baseline training, although caution should be taken because TSS behaves differently in races compared to training (Chapter 6). It seems that combining sRPE with an external load value could be promising in detecting fatigue.

Professional cyclists are highly endurance trained and it could be that multiplying RPE-scores with a high amount of duration, blunts the effect of the RPE-score. In Chapter 6 we showed that sRPE is not affected by exercise intensity and this could mean that training session with a high exercise intensity could be undervalued. Undervaluing sessions with high exercise intensity could be the reason for the fact that Sanders et al. found a weak relationship between the dose measured by sRPE and the fitness response in professional cyclists. A TL measure should be at least associated but preferably related to the response, which could be fitness, fatigue, performance or health. Thus, a weak relationship with fitness seems to be a disadvantage of the sRPE in monitoring training effects in professional cycling.

**luTRIMP**

TL based on HR is frequently used in elite sports, mainly because it is cheap and easy to use. With the introduction of the power meter, professional cycling changed slowly from a HR-based sport to a power-based sport. Although this introduction of the mobile power meter was already introduced in 1986, there are still bicycle disciplines where it is not always possible to measure PO (e.g. mountain biking and cyclocross). Elite athletes participate in those disciplines in the winter period or as part of their season. Therefore, HR based load measurements could be an easy and useful way to fill in the gaps in the periods when PO is missing. In Chapter 5 it is shown that the correlation between luTRIMP and other load values is *very large* or *almost perfect* in training and racing and therefore could be interchangeably used. However, load measures based on HR can be affected by fatigue. Professional cycling is a sport with a high physiological demand and riders compete for up to 10
days without any rest days. It is shown that HR is suppressed in GT$^{22,60}$ and this could result in an underestimated TL measure compared to the actual load the athlete is exposed to. A fatigued cyclist will be exposed to a lower amount of load than planned by the trainer, and thus this could in the eyes of the trainer, lead to a mismatch between the planned load and the load actual exposed to. In a worst-case scenario, the trainer may even increase the TL to get it matched with the training goals and thus putting an already fatigued athlete at risk for negative training adaptations. The reported finding that HR is affected by fatigue can been seen as a disadvantage. However, combining a HR measurement in combination with an external TL measure, which is not affected by fatigue (e.g. kJ spent), could result in a measure to monitor fitness and fatigue in elite cyclists.$^{36}$ Furthermore LuTRIMP showed a stronger dose-response relationship compared to sRPE,$^{16}$ despite the observation that HR is affected by fatigue. As such TL measurements based on HR could be valuable for athletes and coaches.

**TSS**

TSS was introduced by Allen and Coggan as a TL measure based on PO and FTP.$^{23}$ Both Sanders et al.$^{16}$ and Halson$^{43}$ classified TSS as a measure of external load. To my opinion, TSS does not fully meet the definition of either internal or external TL and can be seen as a mix of both. TSS is calculated based on external measurements (i.e. PO and duration). However, adding FTP, although an external measurement, makes that TSS does not match the definitions of either internal or external load measurement. As stated by Wallace:

“External TL is defined as the work completed by an athlete measured independently of his or her internal characteristics”$^{12,13}$

With adding FTP to the calculation of TSS, TSS is not independent any more from internal characteristics. As FTP depends on fitness, fatigue, training background and genetics, we argue that TSS is not independent of internal characteristic and strictly not an external load. Further Drew et al.$^{14}$ defined the differences between internal and external TL as:
“Internal workloads can be described as a measure of perception of effort by the athlete themselves (e.g. rate RPE of heart rate), and external workloads are more typically the quantification of workloads external to the athletes by someone else (e.g. balls bowled in cricket or running distance covered).”

With adding FTP to the calculation of TSS it is not possible anymore for someone else to quantify the TL based on PO because you will need the FTP of the cyclist. Furthermore, Mujika stated that:

“In fact 2 athletes may undertake an identical external training load but experience quite different internal load, depending on their fitness, training background, and genetics.”

TSS is specifically designed by Coggan and Allen\textsuperscript{91} that 100 points accumulated by a professional cyclist is equal to 100 point accumulated by an amateur. When TSS is 100% an external load, a training of 100 TSS points would mean a similar training for a pro or an amateur, but this is not the case, while this is the case for all other external TL measures (throws, distance, kJ spent, etc.). With a training based on external measurements, for example: 100 km at 30 km/h, will result in exactly the same ride for a professional or an amateur. Based on the definitions of internal and external TL mentioned above, TSS cannot be categorized into solely internal or external TL because adding FTP ensures that it does not fall 100% under the definitions of either internal or external TL. This would require adjusting the FTP of a cyclist everyday and calculating the TSS based on this daily influences of fitness, fatigue, training background and genetics. Then TSS could be described as an internal measurement of TL, based on external parameters. However, at this moment it is not possible to adjust FTP on a daily basis and so TSS behaves the same as an external TL measure. This is shown in Chapter 5 were TSS has an almost perfect relationship with external TL measure kJ spent. TSS is not influenced by fatigue, nutrition, hydration or any other parameters if it is not possible to adjust the FTP on a daily basis. The only parameter which is influencing TSS is exercise intensity, with as result a different measure of TSS in training and races. Internal TL is mostly based on the linear relationship between VO\textsubscript{2} consumption and
relative HR or RPE.\textsuperscript{13,113} Other load measures (e.g. iTRIMP\textsuperscript{99} and banTRIMP\textsuperscript{105}) are based on an exponential weighting factor, originated from the classically described exponential rise of blood lactate when exercise intensity exceeds the lactate threshold. Although it is discussable whether TSS is an external or internal measurement of TL, the quadratic relationship with exercise intensity seems to miss a scientific basis. As investigated in Chapter 6 this quadratic relationship with exercise intensity results in different load measurements in training versus racing for TSS in contrast to \(\text{luTRIMP}\), sRPE and kJ spent (Chapters 3, 4 and 5). However, Sanders et al.\textsuperscript{16} showed that TSS has a strong dose-response relationship with aerobic fitness, while \(\text{luTRIMP}\) and sRPE showed the weakest dose-response relationships. This could suggest that the TL measures used in professionally cycling should be more influenced by exercise intensity and less by endurance. Professional road cyclists perform highly endurance-based training exercises and it seems that long but low intensity rides have a limited influence on fatigue. This is strengthened by the results from Chapter 2, showing that climbing performance is affected by short-term fatigue. Surprisingly, fatigue measured by TSS or kJ spent did not have any significant effect on climbing performance but TSS or kJ spent relative to distance did (kJ·km\(^{-1}\) and TSS·km\(^{-1}\)). This could suggest that coaches and sports scientists who want to monitor fatigue in professional cycling should even use a TL measure which reacts stronger to exercise intensity than TSS. Sanders et al.\textsuperscript{36} suggested using TSS in combination with other TL measure to monitor fatigue. However, these suggestions should be taken with caution as it is shown that TSS reacts differently in training and racing compared to the other load measures and thus it seems not valuable to base a fatigue measure on a combination of TSS and one of the other load measures.

**kJ spent**

With the introduction of the mobile power meter, kJ spent is easy to collect in professional cycling. Chapter 5 showed that kJ spent is highly related to internal TL measurements and TSS in professional cycling. This suggests that kJ spent could be a valuable measure for coaches to use as a measure of external TL. kJ spent has some advantages in contrast to other external measures (i.e. PO, distance or duration). A certain amount of kJ spent will always be the same in contrast to distance which is highly dependent on the geography of the
ride. Furthermore, the duration does not give any information about the load, because a 5-hour training ride on low intensity will have a totally different response than a training of 5-hours with high intensity. Therefore, kJ spent in relation with an internal TL measure could be a good tool for the assessment of fatigue or fitness.

kJ spent, sRPE, luTRIMP and TSS have a large to very large correlation with each other. This is valuable information for coaches and sport scientists because this could limit the amount of missing data as load measures can be interchangeably used (Chapter 5). However, this strong correlation means that the TL measures are measuring the same concept, despite that TLs are based on RPE, HR and PO. All the investigated TLs are highly dependent on endurance and thus it could be that the endurance component blunts the effect of RPE, HR and PO. Especially in endurance cycling, were most of the sessions have a high duration, this could be the case. Further, it is questionable if a perfect correlation between internal and external load is desirable. The added value of internal load in comparison to external load is that it is influenced by individual characteristics (e.g. training status, psychological status, health, nutrition, environment and genetics). A perfect correlation suggest that this added value might to be limited. As discussed previously all investigated TL measures have some advantages and disadvantages. Therefore, it would be recommended to always monitor a combination of TLs measures. This limits the risk of missing data or overlooking information. Furthermore, a combination of the TLs measures could be explored to measure fitness and fatigue.

8.3 Future research

Despite some of the novel findings presented in this thesis, there are still some unknowns about the training process in professional cycling. Technological advantages make it easy to collect performance data in professional cycling. However, only PO necessary to compete for the victory in sprint stages is known in both male and female professional cyclists. This thesis added the requirements necessary to compete for the victory in multiple GTs on key mountains and could explain 86% of the variation of the power output on the key mountains. Therefore 14% is still unknown. It could be highly interesting
for sport scientists to investigate the remaining 14% by evaluating more parameters (e.g. nutrition, hydration, environmental and tactical influence). Further, performance values are still unknown for other types of races in both men and women (i.e. (team) TTs and classics) as well as for female GC contenders.

This thesis shows that demands in races and training differ between male and female professional cyclists. It is suggested that also the demands differ between the different types of races. Thus, future research should focus on the load and intensity demands and the performance in the different types of races in men and women professional cyclists and whether different demands also lead to different training strategies in riders with different specialisms. I propose in this thesis a more polarized training approach in professional cycling. This suggestion is based on comparing the training goals with the demands in races. Although, it is still unknown if a more polarized approach would result in better performance in races. Furthermore, high individual differences were seen in the training strategies of professional cyclists (Chapter 4). It is highly interesting to investigate if riders with other training approaches have greater chances of becoming successful.

Chapters 5 and 6 investigated the relationship between various TL used in professional cyclists and high to almost perfect correlations are found. With this strong correlation it can be concluded that those TLs seems to measure the same concept, while coaches want to know which load has the largest associations with positive and negative adaptations. Furthermore, professional cyclists are extremely endurance based trained and the strong relationship between the different TLs could mean that the endurance component blunts the effect of the intensity and thus only endurance is measured. Further Sanders et al. investigated the dose-response relationship of various load measure and found that TSS had the highest dose-responses relationship, although it could be that a novel TL measure with even more influence of intensity could show a stronger dose-response relationship.

In Chapter 7 an attempt is made to investigate which TL ratios have an association with the risk of injuries in professional cycling. Because of the limited number of injuries, it was difficult to conclude which TL ratio is the most useful. Because cycling is a low impact sport the incidence of injuries is lower
compared to other sports. Future research that investigates the different TL ratios should assess even more seasons of data with more injuries. In addition, multiple other factors in professional cycling could increase the risk of injuries (i.e. training conditions, external stressors, the change of materials, bike position and strength training). However, in the study described in Chapter 6 this was not investigated. Therefore, future research should also focus on other factors that increase the risk of injuries. This information could lead to a lower incidence of injuries in amateur and professional cycling.

### 8.4 Practical implications

**Performances during a GT:** This thesis presented PO on the key mountains in multiple GTs from a highly successful GC contender. The performances on the key mountains is influenced by mountain characteristics and short-term fatigue. This knowledge is useful for coaches and sport scientists in the context of talent identification or performance monitoring. Further, in the literature suggestions are made to use performance data (e.g. PO) in the fight against performance enhancing drugs. The presented performance data in combination with identifying factors that influence this performance could be the first step in athlete monitoring systems based on PO.

**Differences between men and women:** In this thesis I presented load and intensity demands for races and training, which are collected in highly trained professional male and female cyclists. The load and intensity demands can be used by coaches and sport scientists for developing a detailed training program for professional male and female cyclists. Further, race regulation and other factors lead to different race demands for male and female professional cyclists. This could also be the case in other sports, and it is recommended to coaches to take those differences into account when working with both male and female athletes.

**Use of Training load measures in professional cycling:** sRPE, luTRIMP, TSS and kJ spent have a large to almost perfect correlation with each other during training, racing and TTs. However, during races, both sRPE and luTRIMP have
a lower relationship with TSS and kJ spent. Especially the suppression of HR with fatigue could lead to an inaccurate TL in highly demanding periods. Coaches and sport scientists working in professional cycling should be aware that TSS reacts differently to exercise intensity compared to other load measures. In addition, the large to almost perfect correlation between the other TL measurements can help coaches and sport scientist to limit the amount of missing data as missing data could be calculate based on this strong correlation. As suggested by Halson et al.\textsuperscript{15} it is important to monitor elite athletes with a combination of TLs instead of using only one TL measure.

Prevention of injury: sRPE is an easy to use, noninvasive and inexpensive tool to monitor athletes with the goal to prevent injuries. A high acute load and a high acute to chronic load ratio could lead to an increased likelihood of injury in professional female cyclists, while a high chronic load protects against injuries. Coaches should be careful with the planning of TL, to avoid peaks in workload ratios.

8.5 Conclusions

The aim of this thesis was to gain a better understanding of the training process in professional male and female cycling. The studies presented in this thesis explored multiple components of the training process in professional male and female cyclists. For the first-time load and intensity demands and the performance of a GC contender are described. This thesis also described the load and intensity demands in professional male and female cycling in racing and training. In addition, the relationship between different internal and external load measures was investigated and the difference between TSS and other load measures was explained by the quadratic relationship with exercise intensity in TSS. Finally, the possible negative response when load and recovery are not in balance is explored in professional female cyclists. The studies presented contribute to a better understanding of the training process in professional male and female cyclists and provide implications for future research. This thesis concludes the following:
• The average performance on the most important mountains in a grand tour is 5.9 W·kg\(^{-1}\) and is influenced by mountain characteristics and short-term fatigue. The load and intensity demands are similar in different grand tours when competing for the General Classification.
• Load and intensity demands differ between men’s and women’s races. Overall volume and absolute load were higher in men’s races while relative intensity and time spent in the high intensity zones (i.e. zones 4 and 5) are higher in women’s races.
• Comparable with the demands in races, absolute loads are higher in men’s training compared to women’s training. Load expressed as TSS and TRIMP are similar in men’s and women’s training. However, relative load and time spent in high-intensity zones are higher during women’s training.
• The TL measures sRPE, luTRIMP, TSS and kJ spent all have a strong correlation with each other in training, road racing and TTs in professional cycling. However, the relationship between TSS and the other TL measures is significantly different during training when comparing to races.
• TSS reacts differently on exercise intensity compared to various other load measures such as luTRIMP, sRPE and kJ spent. Therefore, TSS collected in training is \(~120\%\) higher compared to racing while the difference between training and racing in other TL measures is only \(~70\%\).
• High acute training load and high training load ratios are associated with an increased risk of injuries. While a high chronic training load is protective against injury.


32. Stoggl T, Sperlich B. Polarized training has greater impact on key endurance variables than threshold, high intensity, or high volume training. *Front Physiol.* 2014;5:33.


50. Chicharro JL, Hoyos J, Bandres F, Terrados N, Fernandez B, Lucia A. Thyroid hormone levels during a 3-week professional road cycling competition. *Horm Res.* 2001;56:159-164.


van Erp T, Hoozemans, M., Carl Foster and Jos J. de Koning. The influence of exercise intensity on the association between kJ spent and various training loads in professional cycling. *Int J Sports Physiol Perform.* 2019;In press.


126. Kuylenstierna D. Training Stress Score. *Internal report Chalmers University of Technology (Göteborg).* 2016.


