The puzzle of monitoring training load in winter sports – A hard nut to crack?

Nils Haller*1 2, Tilmann Strepp1, Thomas Stöggl1 3

1 Department of Sport and Exercise Science, Institute of Sport Science, University of Salzburg, Hallein-Rif, Austria
2 Department of Sports Medicine, Rehabilitation and Disease Prevention, Johannes Gutenberg University of Mainz, Mainz, Germany
3 Red Bull Athlete Performance Center, Salzburg, Austria

* nhaller@uni-mainz.de

ABSTRACT

Elite winter sports have become increasingly professionalized and commercialized, often accompanied by congested competition schedules. This has brought the topic of load management, i.e., planning, monitoring, and adaptation, enormous attention in the scientific community. In this review, we summarize the general knowledge from various sports about training load and its effects on injuries and illnesses. After outlining the fundamentals of load monitoring, we present common monitoring tools, i.e., neuromuscular performance tests, heart rate measures, blood-based biomarkers, and questionnaires and discuss their potentials and pitfalls in application. Studies have identified training load-related variables, but also other factors such as travelling, previous injuries, or age contributing to the occurrence of illness and injury. Based on our findings, the use of monitoring tools must be context-specific and long-term, considering statistical aspects as well. Promising study designs and future perspectives are finally highlighted, with the general conclusion that there is still a clear need for research on this topic in general and in the field of winter sports in particular.

Keywords
load management, injury prevention, biomarkers, illness prevention, injury risk, external load, internal load

Citation:
In recent years, elite sports have experienced an increase in professionalization and commercialization, often accompanied by congested competition schedules. These circumstances lead to increased attention to monitoring training and competition loads intended to prevent injury and illness and maximize athlete performance (Halson, 2014; Schwellnus et al., 2016).

Training load and its effects on injury and illness risk have been the subject of numerous research for many years. While the mechanisms and influencing factors contributing to such adverse events in the respective sports often remain elusive, training load is regularly monitored to detect the occurrence of overload and chronic fatigue at an early stage and to adjust training and competition loads. Practitioners often rely on external measures of workload such as global navigational satellite system (GNSS) on the one hand, but also on internal measures such as training diaries and perceived fatigue, heart rate (HR) measurements, or established blood markers to capture the current state of training load, recovery, and fatigue (Halson, 2014).

It should be noted, however, that there is not yet a gold standard approach to load monitoring, as there are large differences between different sports and appropriate load monitoring must always be context-specific. In winter sports this aspect is particularly complex based on the big variation in the demands of the sport and accompanied training load. As an example, ski jumping is highly speed and explosive strength dominated, alpine skiing and snowboarding has additional demands on power endurance and also on the anaerobic/aerobic system while cross-country skiing (XCS) is dominated by endurance capacities, but also substantial speed and strength demands. Nordic combined includes both ends of the spectra.

Here, we summarize basic principles of training load monitoring, the current state of evidence on training load and its effects on illness and injury risk in the field of winter sports as well as popular load monitoring tools in the field of winter sports with their potentials and pitfalls in application.

## Introduction

In recent years, elite sports have experienced an increase in professionalization and commercialization, often accompanied by congested competition schedules. These circumstances lead to increased attention to monitoring training and competition loads intended to prevent injury and illness and maximize athlete performance (Halson, 2014; Schwellnus et al., 2016).

Training load and its effects on injury and illness risk have been the subject of numerous research for many years. While the mechanisms and influencing factors contributing to such adverse events in the respective sports often remain elusive, training load is regularly monitored to detect the occurrence of overload and chronic fatigue at an early stage and to adjust training and competition loads. Practitioners often rely on external measures of workload such as global navigational satellite system (GNSS) on the one hand, but also on internal measures such as training diaries and perceived fatigue, heart rate (HR) measurements, or established blood markers to capture the current state of training load, recovery, and fatigue (Halson, 2014).

It should be noted, however, that there is not yet a gold standard approach to load monitoring, as there are large differences between different sports and appropriate load monitoring must always be context-specific. In winter sports this aspect is particularly complex based on the big variation in the demands of the sport and accompanied training load. As an example, ski jumping is highly speed and explosive strength dominated, alpine skiing and snowboarding has additional demands on power endurance and also on the anaerobic/aerobic system while cross-country skiing (XCS) is dominated by endurance capacities, but also substantial speed and strength demands. Nordic combined includes both ends of the spectra.

Here, we summarize basic principles of training load monitoring, the current state of evidence on training load and its effects on illness and injury risk in the field of winter sports as well as popular load monitoring tools in the field of winter sports with their potentials and pitfalls in application.
expressed as an average post-race ranking, also linking high training load to performance outcomes (Svendsen et al., 2015).

In another prospective study of 91 elite adolescent alpine ski racers, higher training intensity, expressed as ratings of perceived exertion (RPE), was shown to be associated with illness in the same week ($R^2 = 0.12$). In addition, higher training volume in minutes was associated with illness in the following week ($R^2 = 0.16; Hildebrandt et al., 2021$). Of note, the calculated models explained only 12-16% of the variance in the dependent variable of illness, and furthermore, weekly training volume and intensity have not been found to be associated with risk of injury.

Taken together, first evidence of a possible dependence between illness and injury with training load was created. However, it is important to understand the occurrence of both illness and injury as a multifactorial phenomenon. Thus, in addition to the objectively measurable training load and the psychophysiological response, there are external factors (e.g., travelling, season) and internal factors (e.g., perceived stress, pressure) that may further contribute to the occurrence of illness and injury. Figure 1 outlines fixed and modifiable illness and injury risk factors in winter sports.

**Training load monitoring and its challenges in winter sports**

Training load monitoring is performed to assess performance capabilities and to determine how athletes are responding to an objective training load. Therefore, training load typically incorporates aspects of both the objective physical workload, i.e., the external load – and the psychophysiological response, i.e., the internal load (Impellizzeri et al., 2019). In a recent consensus statement on load monitoring, the authors (Bourdon et al., 2017) summarized the numerous monitoring tools from GNSS tracking of the external load to biomarkers, questionnaires, and neuromuscular performance as estimates for the internal load. This leads to the question of which instrument should ideally be used. In many intermittent sports such as soccer, basketball or field hockey, the external load can be well captured e.g., by GNSS or local positioning system (LPS) tracking with measured variables such as total distance, sprints, intensive runs. In winter sports, the requirements between the individual winter sports range from strongly endurance-oriented to a mixture of maximum strength, coordinative, but also mental requirements. Similarly, the internal load measures must also be analyzed in a highly context-related manner. For example, HR in winter endurance sports may be a good measure of internal load, whereas in ski jumping there is insufficient evidence that HR reflects the athlete's effort. Therefore, a consensus should be reached among experts on the important variables of external and internal load in the different winter sports. The same applies for the measure of external load based on position and speed of motion with challenges based on the undulating terrain, snow, and ski/wax conditions in XCS and alpine skiing. The measure of power output in the various winter-sport disciplines might be a future solution with miniaturized and wireless sensor technologies (e.g., force sensor instrumented equipment, inertial measurement systems, or a combination of both).

**Measures of internal load and fatigue**

In the following subsections, several instruments researched in winter sports for the assessment of internal load or fatigue status, as well as their advantages and disadvantages, are presented. Ideally, load monitoring tools should meet the following criteria; i) sensitivity, ii) knowledge of the physiological background (such as origin, time-course of e.g., biomarkers), iii) not too invasive, iv) not too expensive, v) no influence on the training routine, vi) repeatedly measurable vii) showing early signs of overload (Meeusen et al., 2013; Schwellnus et al., 2016).

**Questionnaires**

Questionnaires, such as the Daily Analyses of Life Demands of Athletes or the Profile of Mood States are reliable tools reflecting changes in the training load.
Moreover, they can obtain different constructs such as training load, but also recovery, mood, stress, and further constructs. A traditional tool for monitoring the internal training load is the RPE on a scale from 0 to 10 (Foster et al., 2001), which is sometimes multiplied by total training time (Impellizzeri et al., 2004). While RPE can also be used to differentiate between whole body RPE and leg RPE depending on the context (Stöggl et al., 2016), the RPE also serves as a variable to determine training intensity distribution (Seiler & Kjerland, 2006). Besides this important easy-to-use tool to monitor the internal training load, further constructs may be considered. One of them is sleep quality and quantity. Sleep disturbances are common among elite athletes, and strategies to improve their sleep are often lacking (Erlacher et al., 2011). A recent study with 56 junior endurance athletes from high schools specialized for winter endurance sport conducted daily monitoring of sleep (with device and questionnaire), mental strain, and training load. The authors concluded that all constructs have a reciprocal relationship; that is e.g., increased mental strain is related to reduced sleep time, increases in sleep time are related to reduced training load and mental strain as well as training load are associated with decreased rapid eye movement measured by device (Hrozanova et al., 2020).

**Figure 1** Possible illness and injury risk factors in winter sports. Modifiable risk factors related to training load in red (Extracted from Bere et al., 2014; Engebretsen et al., 2010; Hildebrandt et al., 2021; Malone et al., 2018; Svendsen et al., 2015, 2016).
In practice, an individualized approach to identifying “bad sleepers” (zum Berge et al., 2021, p. 96) may be a promising approach to maximize performance.

These findings should serve as examples how versatile questionnaires can reflect different parts of the training load and its immediate consequences on sleep or mental aspects. Questionnaires or training diaries should thus be included in routine monitoring. However, practitioners must also be aware that there is a risk of giving socially desirable answers and over- or underestimating the training load. The assessment of sleep quality and quantity is also highly subjective and may be prone to error. Therefore, modern developments in sleep monitoring (e.g., HR monitoring, bed-integrated approaches such as ballistocardiographic (BCG) sleep-tracking devices (Vesterinen et al., 2020) special rings or wristbands) could be valuable add-on tools to combine subjective and objective tools, which is highly recommended.

**Neuromuscular performance testing**

Neuromuscular performance testing is usually conducted before the onset of exercise using squat jump (SJ) or counter movement jump (CMJ). The highest or average jump height over multiple repetitions is discussed to reflect aspects of athlete readiness and fatigue (Claudino et al., 2017). In elite junior alpine skiers, during an 11-day high-intensity training (HIIT) shock cycle, CMJ and SJ maximal power output was impaired in the intervention group, suggesting a chronic decrease of neuromuscular performance, even though endurance performance was shown to increase at that time point (Breil et al., 2010). Interestingly, a positive relationship was found between CMJ height and VO$_{2\text{max}}$ after one year of training in young XCS (Mishica et al., 2023). In contrast, during 3 weeks of similar training, no significant changes in jump height were observed (Mishica et al., 2022).

However, key questions remain to be addressed. First, it remains elusive, whether the best attempt or the average value of several jumps should be considered a meaningful result. In addition, a single jump results in many variables that can be analyzed. Further CMJ variables related to movement quality could also be impaired in fatigued athletes. Therefore, a set of jump variables including both output and movement strategy should be considered in the analysis in future studies (Gathercole et al., 2015; Haller et al., 2022).

**Biomarkers**

Biomarkers, i.e. indicators of normal biological processes, pathogenic processes, or responses to an exposure (FDA-NIH Biomarker Working Group, 2016), measured in the blood (or other fluids, e.g., urine or saliva), are capable of delving deeper into the physiology of the athlete. There are several rationales for collecting blood biomarkers, such as measuring acute exercise responses, recovery time courses, and general health (Haller, Behringer, et al., 2023). Established biomarkers such as lactate or creatine kinase (CK) are routinely measured during and/or after exercise to estimate the acute exercise intensity or the recovery status (Brancaccio et al., 2007; Dickhuth et al., 1999). Cortisol and testosterone were even discussed as promising markers for the overtrained state. However, none of these parameters have been extensively studied (Cadegiani & Kater, 2017), especially in the field of winter sports. One study in 16-year-old XCS and biathlon athletes over a seven-week period during training and competition showed that cortisol levels were elevated during the competitive season, with associations to recovery (Mishica et al., 2021). Significant changes in immunological parameters of XCS were detected over two months during the competitive season (Mueller et al., 2001), highlighting the added value to monitor biomarkers of inflammation in regular load monitoring.

Regarding the acute exercise response, immunoenocrine and metabolic responses were measured after a 50 km (men) or 30 km (women) XCS race, with marked increases in biomarkers such as granulocytes and epinephrine (Rønsen et al., 2004). While changes were highly interindividual, sex differences further complicated the interpretation of concentration changes. For example, serum testosterone remained unchanged in male athletes, but doubled in females.
The findings from previous studies and from biomarker research in general is advised to be applied to winter sports. As mentioned above, biomarker concentrations and changes due to training load are highly intra- and inter-individual and often multifactorial. This complicates the detection of meaningful or clinically relevant changes (i.e., changes that differ from the measurement error and may require attention; Swinton et al., 2018). Longitudinal observation is therefore required to establish baseline values for the selected biomarkers using appropriate statistical methods. Importantly, it should be noted that a single biomarker basically reflects a limited physiological domain, such as muscle damage. Depending on the sport, complex physiological response patterns may be related not only to muscle damage but also to inflammation or an immune response. Therefore, more than one marker is needed to fully capture the physiological response of exercise. Finally, the measurement of biomarkers requires a team approach in terms of test standards, hygiene, sampling, analysis, and evaluation, and therefore requires time and effort (Haller, Behringer, et al., 2023).

Due to the limitations of established biomarkers such as CK or lactate, emerging biomarkers as well as novel bioanalytical methods such as proteome analyses, which determine a large number of markers simultaneously, are being explored. However, these are still far away from routine application (Haller, Reichel, et al., 2023).

**Measures of heart rate**

Heart rate is a widely adopted, non-invasive and inexpensive method across many sports to quantify the internal load of athletes (Achten & Jeukendrup, 2003; Jeukendrup & Van Diemen, 1998). The application is versatile and ranges from sleeping, resting, and exercise conditions recorded with different types of devices (e.g., chest strap, electrocardiography, BCG).

For example, increased sleeping HR was observed in overtrained individuals after two weeks of intensified training (Jeukendrup et al., 1992). It has been argued that sleeping HR is a more reliable measure compared with resting HR because it is less likely to be influenced by confounding variables (Jeukendrup & Van Diemen, 1998). A recent study compared sleeping HR determined by BCG and morning HR measured with a chest strap, showed no significant difference in the method in young well-trained XCS (Mishica et al., 2022). Resting HR has been discussed to provide insight into the current fatigue state, well-being, and performance readiness of athletes (Bosquet et al., 2008; Buchheit, 2014; Djaoui et al., 2017). Typically, an increased resting HR during regular training periods is a sign for accumulated fatigue, while decreased resting HR is linked to positive coping with training. In contrast, reduced resting HR was observed in an overtrained XCS, which illustrates, that changes in resting HR have to be contextualized with the external load and the individual training history of the athlete (Hedelin et al., 2000).

Due to the somewhat linear relationship between oxygen uptake and intensity, HR can also be used to reflect exercise intensity (Arts & Kuipers, 1994; Gilman, 1996). However, its application in field-based winter sports is limited. For instance, XCS includes uphill sections with high anaerobic energy contribution where HR may overestimate oxygen consumption (Crisafulli et al., 2006). In addition, the latency in HR response to a change in exercise intensity further limits its use in hilly terrain (Born et al., 2017). Phenomena like cardiac drift, or systematic reduction in HR especially in the long distance XCS races can further complicate the analysis (Stöggl et al., 2020). Finally, decreased HR during maximal exercise has been associated with overreached individuals (Achten & Jeukendrup, 2003; Costill et al., 1988). Seven days after a HIIT shock microcycle (15 HIIT sessions in 11 days), a significant decrease in maximum HR of 3 beats per minute (bpm) was observed in elite alpine skiers, while maximum oxygen uptake (VO2max) improved by 6% and peak power by 5.5% (Breil et al., 2010). It remains elusive whether such a state is desirable for certain adaptation processes or if such suppression of maximal HR affects the training effect at all. Nevertheless, day-to-day variability (2-6 bpm), cardiac drift (dehydra-
tion & heat stress), environment (temperature & altitude) or overtraining may impact HR, leading to possible over- or underestimation of training load and should therefore be considered with caution (Achten & Jeukendrup, 2003; Bagger et al., 2003; Lambert et al., 1998).

A promising method for tracking training adaptions and recovery is the monitoring of the sympathetic and parasympathetic activity of the autonomic nervous system through HR variability (HRV; i.e., the beat-to-beat variation in the heart; Dong, 2016; Plews et al., 2013). A decrease in HRV is commonly associated with non-functional overreaching and/or negative adaptation to training (Bosquet et al., 2008; Hynynen et al., 2006), while an increase is linked to the recovered state, improvements in fitness and exercise performance (Buchheit, 2014). There is also evidence that HRV may be reduced before stressful situations, anxiety, or complex decisions, potentially important for competition (Dong, 2016; Miu et al., 2009).

In eight young winter athletes, during a seven-week period with nocturnal HRV measurement using BCG, physical workload during training and competition phases was well reflected by HRV (Mishica et al., 2021). In particular, the decrease in HRV correlated with a lower amount of intense training/training load. HRV can therefore help to identify the relationship between external and internal loads. Schmitt, Willis, et al. (2018) subjected elite XCS and Nordic combined athletes to a 15-day HRV guided training in live high-train low conditions to decide on daily individualized adjustments of the training load. The guided and non-guided groups improved their VO2max by 3.8% (p = 0.02) and 3% (p = 0.08), respectively, with only the guided group improving their roller ski performance by -2.7% (p = 0.05) 21 days after the intervention. However, during regular monitoring over 5 years in 57 elite Nordic-skiers consisting of XCS, biathlon, and Nordic combined athletes, no causal relationship between training load/intensity and HRV fatigue patterns was observed (Schmitt, Regnard, et al., 2018).

Taken together, HRV in winter sports has potential to optimize training load management to enhance performance. Of note, longitudinal HRV monitoring is crucial to determine the optimal value for each athlete to detect meaningful changes.

Future perspectives
The research on load monitoring is evolving. In case of biomarkers, this is true for e.g., point-of-care devices to rapidly measure novel biomarkers such as cytokines (Haller, Reichel, et al., 2023). For other tools, Stögg & Born (2021) equipped one former elite XCS with four Near Infrared Spectroscopy devices, a GNSS, and a HR monitor during the Vasaloppet long-distance XCS race showing that Triceps brachii muscle oxygen saturation may serve as an alternative measure for exercise intensity (Stögg & Born, 2021). In another study, body temperature in XCS was measured by infrared thermography, and some correlations between average body temperature after exercise and other physiological, biochemical, and morphological parameters were observed. However, this was only a pilot study with small sample size in a non-ski specific setting (Drzazga et al., 2018).

As technology evolves, practitioners are faced with a large amount of data especially when using several monitoring tools. Machine learning (ML) could help them with data analysis and decision making for load management. ML approaches have been shown to predict injury (in sports such as soccer, handball, or basketball), with sleep quality, genetic variables, and external load parameters as key variables (Van Eetvelde et al., 2021). However, the quality of studies has been low to moderate, and to our knowledge, no such approach has been taken in winter sports thus far. However, predicting athletes’ responses to training could help winter sports practitioners manage upcoming training loads (Bartlett et al., 2017).

Study designs needed to develop reliable approaches for load monitoring
Load monitoring in winter sports is still in its infancy with many instruments not extensively researched in terms of their reliability, predictivity or their associ-
ation to injury or illness risk. Publications often conclude with sentences such as "monitoring [...] training load between weeks may decrease incidents of illness in-season" (Hildebrandt et al., 2021, p. 230) which are currently rather speculative. Apart from a possible link between workload and adverse events in the context of a multifactorial genesis, the usefulness of monitoring tools with regard to injury and illness prevention has hardly been demonstrated so far. In an innovative pilot study (Sperlich et al., 2016), training load was adjusted in response to monitoring tools exceeding a predefined range (e.g., CK values that fell outside defined reference ranges). They found an increase in performance in athletes without detecting injuries. Although this study lacked a control group, these types of studies could further investigate the relationship between load monitoring and the occurrence of adverse events.

In addition, longitudinal research will further advance the usefulness of monitoring tools. Our research group recently conducted a one-month intervention study using a 7-day HIIT shock cycle as a model to examine a range of monitoring tools for exercise load, sleep, muscle fatigue, inflammation, neuromuscular performance, and cognitive function. The study included a baseline phase, a 7-day overload period, and 14 days of recovery (Stöggl et al., 2022). This study design allows for repeated measures to capture both the acute as well as the chronic changes in monitoring tools. In addition, the variety of instruments covering different domains related to training load, fatigue, and recovery allows the choreography of physiological responses, e.g., using advanced multilevel analyses to investigate the interdependencies between variables.

**Practical applications**

All instruments have strengths and weaknesses and do not meet all the above defined criteria. It is important to note that the ideal tool can only be determined depending on the objective as well as the sport-specific context. Because many monitoring approaches in winter sports remain under-researched or even unpublished, it could be worthwhile to look at successful coaches as they are experienced to see and learn how their athletes respond to training load. From what they see and learn from the athlete on a day-to-day basis often determines whether training goes as planned or should be modified. Picking the brains of the top coaches could thus be included in the process of determining appropriate monitoring tools. If appropriate instruments are selected, they will need to be measured frequently to i) establish an individual baseline and ii) not miss important time points where meaningful changes occur. Sperlich & Holmberg (2017) highlighted the aspect of continuous measurement by pointing out that activities outside of training significantly affect physiological adaptations. In elite athletes, about 80% of the day is spent on activities such as recovery, e.g., massages, physiotherapy, medical treatments, meals, and activities of daily living (Sperlich & Holmberg, 2017). At the same time, however, the burden on athletes must be kept low. For example, it is well known that athletes are e.g., reluctant to undergo repeated blood sampling (Carling et al., 2018). Thus, depending on the method used, monitoring must be close-meshed but not intrusive to ensure that athletes comply and training as well as recovery processes are not negatively affected.

It is also important to determine the threshold for meaningful changes (Halson, 2014) to be able to intervene as a practitioner, e.g., adjusting training load. Mathematical models, such as the smallest worthwhile change (Hopkins, 2000), the acute to chronic workload ratio (Gabbett, 2016), the ratio between external and internal load (Bourdon et al., 2017) and more recently artificial intelligence based learning, can help practitioners make informed decisions and identify meaningful changes. While practitioners should use monitoring tools with confidence and use them as a help for decision-making, significant changes in a monitoring tool over time are not necessarily a sign to immediately adjust the training load, as these changes may have multifactorial nature. For example, a change in HR during exercise may be a sign of positive adaptation, but it may also be a sign of the onset of overload. Therefore, meaningful changes should be understood...
as a sign to consult with the athlete about the cause of these changes.

**Conclusion**

The area of load monitoring is on the rise in sports in general and winter sports in particular. Many interesting tools have been presented here, although all tools have strengths and weaknesses. Longitudinal research and intervention studies are needed to assess the reliability and usefulness of markers and their ability to reflect acute and chronic training load and fatigue as well as the association with illness and injury. We believe that modern statistical methods should be applied to uncover interactions and relationships between variables that cannot be captured by classical statistics.

**References**


N. Haller et al. The puzzle of monitoring training load in winter sports – A hard nut to crack?


Acknowledgements

Funding

The authors have no funding or support to report.

Competing interests

The authors have declared that no competing interests exist.

Data availability statement

All relevant data are within the paper.