Word, table and Figure counts

Entire document:

Abstract:

Main text:

Number of tables in manuscript:

Number of Figures in manuscript:

Number of references:

Towards a complex systems approach in sports injury research: Simulating running-related injury development with Agent-Based Modelling

Adam Hulme[1], Jason Thompson[2], Rasmus O. Nielsen[3], Gemma J.M. Read[1], Paul M. Salmon[1]

Corresponding author

Dr Adam Hulme

E-mail: ahulme@usc.edu.au

[1]Centre for Human Factors and Sociotechnical Systems

Faculty of Arts, Business and Law

University of the Sunshine Coast

Sippy Downs

Queensland 4556

Australia

[2]Transport, Health and Urban Design (THUD) Research Hub

Faculty of Architecture, Building and Planning

Melbourne School of Design

The University of Melbourne

Melbourne

Victoria 3010

Australia

[3]RunSafe Research Group

Section for Sports Science

Department of Public Health

Aarhus University

Dalgas Avenue 4

Aarhus, 8000

Denmark

ABSTRACT

There have been recent calls for the application of the complex systems approach in sports injury research. However, beyond theoretical description and static models of complexity, little progress has been made towards formalising this approach in way that is practical and meaningful to sports injury scientists and clinicians. In this paper, we used a computational modelling method known as Agent-Based Modelling (ABM), and applied it in the context of distance running-related injury (RRI). The primary aim was to demonstrate the method and simulate the dynamic relationship between the absolute weekly running distance and RRI, as well as the relative change to weekly running distance and RRI, through the manipulation of various athlete management tools'. The ABM was developed based on sports injury and RRI causal theory, and incorporates evidence gathered from studies using the acute:chronic workload ratio (ACWR), an approach that calculates changes in training load. The findings show that building weekly running distances over time, even within the reported ACWR 'sweet spot', will eventually result in RRI as athletes reach and surpass their individual physical workload limits. Specifically, introducing training-related error into the simulation and the modelling of a 'hard ceiling' dynamic resulted in a higher RRI incidence proportion across the population at higher absolute workloads. The implications of this proof-of-concept study extend beyond the presented model. In particular, the ability to dynamically simulate known sports injury mechanisms offers a practical starting point to further examine and/or apply more sophisticated computational models that can account for the complex nature of sports injury aetiology. Alongside traditional forms of scientific inquiry, the use of computational modelling approaches should now occupy a methodological and analytical space in sports injury research.

Keywords: Agent-Based Modelling, Complex Systems, Sports Injury, Distance Running

INTRODUCTION

In the sports injury aetiology and prevention research field, the use of the 'complex systems approach' has been recently promoted[1]. Inspired by previous work that questioned the routine application of reductionist scientific methodologies and statistical techniques[2-4], Bittencourt and colleagues[1] argued for a new causal approach. This approach recognises that sports injury is a 'complex emergent phenomenon', resulting from the interactions among different factors (i.e. a web of determinants), which may produce regularities (i.e. a risk profile), that are antecedent to the emerging pattern (i.e. sports injury). In brief, the theoretical assumptions underpinning the complex systems approach can be traced back to general systems theory[5], which identifies several characteristics of complexity as a general philosophical precept. These include, but are not limited to, adaptation and learning, tight coupling, causal feedback, non-linear relations, sensitivity on initial conditions, threshold effects, stochasticity, and historical dependency. Those characteristics have featured across multiple discussions in the sports injury scientific literature[1 4 6 7], however no study has yet applied a recognised method which has the capability to dynamically simulate and better understand complex systems causal patterns and processes. One computational modelling method that shows early promise, and has been suggested as a suitable approach for sports injury research[1 2 4 8], is Agent-Based Modelling (ABM).

In this paper, we develop a first-of-its-kind ABM in the field of sports science, and apply it in the context of distance running-related injury (RRI). Distance running is considered a pertinent example to use from a technical point of view given that the main participatory-related exposure can be readily defined. There are three aspects to the model which underpin the concept, design, and operation: (i) its development is based on the complex systems approach[1], and so investigating the application of a novel complex systems method in sports injury research represents an important line of inquiry; (ii) the ABM is informed by contemporary sports injury and RRI causal theory[9-11]; and, (iii) it incorporates the acute:chronic workload ratio (ACWR)[12 13], an evidence-based approach that calculates relative changes in training load. The primary aim of the ABM is to simulate the dynamic relationship between the absolute weekly running distance and RRI, as well as the relative change to weekly running distance and RRI, through the manipulation of various model parameters (see section titled, 'athlete management tools').

From a methodological standpoint, using ABM to simulate the relationship between workload and sports injury development is a considerable step forward in terms of complexity science and the systems thinking sports injury literature[1 4 6 7 14 15]. Indeed, during the last decade, there have been a series of articles that share a number of progressive commonalities. Initially, Quatman et al[2] proposed a conceptual-methodological framework encompassing the integration of *in vivo, in vitro,* and *in silico* techniques to better understand the development of anterior cruciate ligament injury. In particular, the authors[2] stated that the greatest advances in sports injury research are likely to come from a new methodological paradigm that allows us to think, theorise, and locate appropriate applications that consider the nature of the complex relationships among different exposures. Shortly afterwards, Mendiguchia et al[3] argued for the same paradigm shift, advocating that sports injury research was required to move beyond the process of wanting to isolate risk factors, to a conceptual model encompassing 'dynamic simulations' and the possibility to 'modify different parameters'[3]. More recently, Hulme and Finch[4] and Bittencourt et al[1] have suggested the use of systems dynamics modelling and ABM for the explication and testing of theoretical causal assumptions in relation to injury development, as well as for the simulation of complex sports injury aetiologic mechanism(s). Further systems-based work has since applied a method from the human factors and ergonomics domain[16], and developed a more holistic, 'complex systems model' of RRI development and prevention[14 15]. Notwithstanding the evolution of systems thinking applications in sports injury research, most scholarly contributions have been descriptive in nature[1-4 6 7], or have involved the development of static frameworks and models[14 15]. As such, to advance this body of work, it is necessary to apply a computational modelling approach which can simulate dynamic behaviours within complex sports systems, and/or understand how systems change over time.[14 15]

With regard to scientific theory and clinical practice, the use of ABM has the potential to generate new findings and insights about sports injury aetiology and prevention, which can then be used to support clinical decision making. Healthcare practitioners rely upon a wide range of study designs and different forms of evidence in which to prescribe the most efficacious therapeutic or preventive interventions to athletes[17-19]. For that reason, there is a need to investigate how, as a proven complementary method to routine epidemiologic inquiry[20-23], ABM can dynamically simulate known mechanisms of sports injury (i.e. workload and RRI), so that it is possible to develop more sophisticated and clinically-relevant complex systems models moving forwards. In taking the next formal step, this study represents the transition away from theoretical description and static modelling approaches[1-4 6 7 14 15], and examines the feasibility of computational modelling for studying the complex and dynamic nature of sports injury moving forwards. Therefore, the purpose of this study is to introduce computational modelling to sports injury research, and demonstrate the potential utility of ABM as a viable method for studying complex injury dynamics in future theoretical and practical applications.

METHODS

Agent-Based Modelling

As a computational method, ABM simulates the actions and interactions of heterogeneous, autonomous 'agents', to assess the effects of their behaviour on the system as a whole[24 25]. Agents in an ABM can constitute any self-contained and goal directed entity, including but not limited to, molecules, cells, pathogens, people (e.g. athletes, runners, sports teams), animals, automated vehicles, organisations, and/or entire synthetic populations[26 27]. In the case that the agents are representative of individual persons, the model operator can assign demographic and lifestyle-related characteristics such as age, sex, diet, medical history, and injury susceptibility, as well as cognitive rules pertaining to memory, personality, behaviour, and/or intelligence[28]. This means that agents can learn over time based on past experiences, update their internal states, adapt to changing environmental circumstances, and demonstrate any other characteristic or behaviour that has been explicitly defined. Based on its 'ground-up' modelling approach, ABM can be used to explain how populations self-organise, and create patterns of global behaviour that are not predictable or programmed into each agent type *a priori*. For this reason, ABM is a powerful tool when wanting to ascertain, by what mechanism(s), the collective behaviour among individual agents gives rise to emergent-level phenomenon (e.g. sports injury). The model itself simulates the passage of time in discrete 'time steps', each step of which can correspond to either seconds, minutes, days, weeks, years, or decades depending on the purpose of study.

Many different health-related contexts have applied ABM. A notable example is the Global-Scale Agent-Based Model, which simulated 6.5 billion persons, and explored how various behaviours and contact points shaped the transmission rate and distribution of the H1N1 swine flu virus[26 27]. Other studies have integrated ABM with Geographic Information Systems science to improve comprehension of how the measles disease propagates through an urban environment[29]. In the non-communicable health context, ABM has been used for multiple purposes[30], including the evaluation of policy-level and environmental intervention strategies for improving diet and promoting exercise[31-33]. Specifically, Yang and co-workers[31] used ABM to examine the impact of certain policies aimed to change population-level attitudes towards walking among individuals from different socioeconomic backgrounds. In the medical and healthcare context, ABM has emulated a real-world lifestyle modification program for individuals with diabetes, and estimated the morbidity and economic outcomes associated with the modification of certain parameters (e.g. pharmacologic delivery options) over a 30-year period[34]. The latter model was instantiated with empirical data, and used sensitivity analysis to ascertain the degree of model error. Since initial applications in the mid-1990s, the use of ABM has continued to gain popularity in parallel with the evolution of information technology and computing power[35]. For further information pertaining to the origins, purpose, and general use of ABM, the reader is referred to other more comprehensive articles[24 25 28 35].

The distance running Agent-Based Model

The distance running ABM was constructed using the NetLogo toolkit (version 6.0.1), a cross-platform, open-source, programmable modelling environment for simulating natural and social phenomena (https://ccl.northwestern.edu/netlogo/)[36 37]. The simulation environment was representative of a track and field overlay with dimensions of 70 x 30 patches (arbitrarily scaled distance units) (Electronic Supplementary Materiel Figure 1). To guide the reader through the different stages of ABM development, the following four phases are described: (i) defining the personal characteristics of the synthetic agent population (hereby referred to as 'runners', or the 'running population'); (ii) initialising the model and establishing baseline procedures; (iii) implementing four distinct 'athlete management tools' (i.e. these tools drive the dynamics of the model); and, (iv) establishing the conditions surrounding the execution of the simulation procedure itself.

Phase one: running population characteristics

The distance running ABM contained 1000 runners. This number of runners was chosen for the purpose of being able to capture aggregate, population-level dynamics and patterns. Each runner possessed personal characteristics that could affect their physical capacity to tolerate an applied external running workload, defined as the number of kilometres (km) undertaken in any given training week (p/w). These characteristics included body mass index (BMI), biomechanics, footwear, sleep, diet, recovery, and genetics, and were selected based on a recent framework of RRI aetiology[10]. To support a comparison of those factors across runners, the relative 'quality' of each characteristic was standardised and drawn from a random-normal distribution with a mean of zero (e.g. an 'average' diet or sleep quality), with a standard deviation of 0.3 (i.e. this maintained most runners between a range of +1 and -1).

An additional characteristic that each runner possessed was a maximum workload potential (MWP) state. In accordance with contemporary sports injury and RRI causal theory[9-11], surpassing the MWP state was equivalent to the absolute external running workload exceeding a specific musculoskeletal structure's physical capacity. Safely reaching the MWP state without surpassing it assumes a perfect environment, training, and management regimen. For each runner in the ABM, their initial MWP state was set to a random-normal mean of 65.0km p/w, with a standard deviation of 10.0km p/w. This produced a population-based MWP distribution that acknowledged not all runners had an equal upper workload limit. It is worth noting that the selection of runners' characteristics and the MWP state values are not necessarily integral to the operation of the simulation or the validity of its outputs. Rather, this proof-of-concept approach was focussed on demonstrating how ABM can be both programmed and used to simulate the relationship between workload and RRI risk – and by extension – overall population-level athletic performance.

The acute:chronic workload ratio

Central to the distance running ABM is the ACWR[12 13]. As a means of facilitating sports performance optimisation, the ACWR can be used to guide the prescription of future workloads. There are two components to this metric: (i) the 'acute' phase, which represents the training load undertaken in the most recent week (i.e. a 1-week block); and, (ii) the 'chronic' phase, which signifies the average training load undertaken in the month prior (i.e. a 4-week block)[12]. Calculating the ACWR involves dividing the acute phase (e.g. 60.0km of running), by the chronic average (e.g. 50.0km), giving in this case a ratio of 1.2 (i.e. 20.0% workload increase).

The ACWR is theoretically driven and practically appealing. Well-developed physical qualities and musculoskeletal adaptations produced during chronic training phases will build athletic resilience and protect against injury[12 13]. Gradually increasing workloads, and closely tracking week-to-week changes to training regimens, is more important than the absolute applied workload exposure at any given time[38 39]. Prospective epidemiological investigations have found that when the acute training load is equal to, or less than, the chronic phase (i.e. ACWR ratio ≤1.0), the risk of non-contact, soft-tissue injury is significantly lower than ratios of ≥1.5[40]. Although further research is yet to be conducted to strengthen existing evidence e.g.[41-46], an ACWR between 0.8 and 1.3 has been coined the 'training sweet spot', whereas a ratio between 1.3 and 1.4, and ≥1.5, represents a moderate and high risk injury zone, respectively[12 47].

In response to the growing interest around the ACWR, concerns have been raised about the use of rolling averages to assess workload and sports-related injury risk[48-50]. The two main limitations with the traditional ACWR calculation are: (i) averages fail to account for variation over time such that day-to-day patterns and 'spikes' in the applied workload are smoothed; and, (ii) rolling averages neglect the decaying nature of stimuli over time[48]. As such, a non-linear training model that places increasing weighting on the daily workloads undertaken towards the end of a chronic training phase has been proposed[50]. The exponentially-weighted acute:chronic workload ratio (EW-ACWR) was found to be significantly more sensitive than the traditional ACWR at identifying injury likelihoods at upper training load ratio ranges (i.e. ≥1.5)[49]. Along with the traditional ACWR, the EW-ACWR was incorporated into the distance running ABM as an option for calculating the relative variation in runners' workload.

Phase two: model initialisation

At the start of the simulation, each runner was assigned a standard running history spanning the previous 28-day period to standardise the model. This history allocated a total of 20.0km p/w in each 7-day block preceding each day in the prior 28-day period. Therefore, at ABM initialisation, each runner had 20.0km p/w history of running in the previous training week, and had a rolling average of 20.0km p/w for the past month. This produced both an initial ACWR and an EW-ACWR of 1.0 (i.e. each runner had a consistent workload in relation to a previously recorded workload over the past 28 days as calculated under each regimen).

Phase three: athlete management tools

The distance running ABM incorporated four athlete management tools that were manipulable by the model operator. The first of these tools is the 'ramp-up rate slider', which dictated the rate at which runners applied and increased workload over time. Specifically, the goal of each runner in the system was to maximise the absolute distance they were able to run per week without sustaining RRI. That is, safely reach the MWP state and remain there. To achieve this, the simulation started with runners' gradually increasing their weekly kilometres at a rate determined by the user-defined ramp-up rate. The lower the ramp-up rate, the longer the time frame before the running population reached a MWP state. Conversely, higher ramp-up rates resulted in runners rapidly ascending to their upper workload limits. For the purposes of experimentation, runners' workloads were increased within the reported ACWR sweet spot of between 5.0% and 30.0%, in increments of 5.0%[12]. This resulted in a total of six possible ramp-up rate conditions.

The second athlete management tool was an ability to approximate a runner's individual adherence, misrepresentation, or miscalculation of the advice provided by, for example, a coach or healthcare professional as to how much training should be undertaken per week. The 'random variation slider' introduced noise into each runner's planned workload, adjusting the ramp-up rate by a mean of 0.0%, but with increasing standard deviations of 0.0% (i.e. perfect training adherence), 1.0% (i.e. moderate training adherence), or 2.5% (i.e. poor training adherence). The formula for the calculation of workload in the current week is shown in Equation 1, where *cw* = workload in the current week, *pw* = workload in the previous week, *r* = ramp-up rate, and *error* = random variation.

Equation 1. $$cw = pw * (1 + (r + error))$$

The third athlete management tool that was manipulated by the model operator was the ability to adjust the way in which the ACWR was calculated. Although differences between the traditional ACWR and the EW-ACWR metrics correspond to approaches that calculate changes in workload, the distance running ABM was constructed so that altering the estimation of the ratio (i.e. either non-weighted or weighted) could affect RRI risk. Therefore, the risk of RRI in each week was based on either the ACWR or EW-ACWR calculation, of which both dynamically responded to the user-defined ramp-up rate and random variation condition. Accordingly, if a given runner's calculated workload ratio was ≥1.1, then the likelihood of RRI was proportional to the cubed value of their allocated ACWR or EW-ACWR condition. This produced an exponentially increasing risk of RRI that approximated the observed likelihood of sports injury development as found in empirical studies (Figure 1)[12]. To provide a visual indication to the model operator of the health of the running population at any given time, runners who incurred a

RRI doubled in size, turned red in colour, and were transferred to the centre of the simulation environment. Upon sustaining RRI, a given runner's workload dropped to 5.0km per week.

<<<<< Insert Figure 1 about here >>>>>

The fourth and final athlete management tool under manipulation was a binary condition relating to the runner's individual MWP state. In the latter condition, whereby an individuals' MWP was unknown (i.e. 'off'), the calculation of runners' future training was based on the workload in the most current week, multiplied by the ramp-up rate, and adjusted for random variation (Equation 1). A second condition was constructed whereby runners' workloads were further adjusted based on how close the current workload was to their MWP state. This calculation recognised that the running population had a randomly distributed MWP that was guaranteed to be reached under conditions of continuous, compounding growth (i.e. MWP state 'on'). The formula for this calculation is shown in Equation 2.

Equation 2.

$$cw = pw * (1 + \left(\left(r * \left(1 - \left(\frac{pw}{mwp}\right)\right)\right) + error\right))$$

Phase four: establishing the conditions of the simulation

The set combination of the six ramp-up rates, three random variation conditions, two ACWR calculations, and two MWP states produced a 72-condition matrix. Given stochastic elements within features of the ABM, modelling under the 72 different possible conditions was repeated 10 times for 1000 model time steps, or days (~143 weeks). This produced a total computational model encompassing 720,000 individual simulated runners monitored over a total of 720,000 days (~102, 800 weeks). Upon completion of the simulation, data were exported from the NetLogo[36 37] software into spreadsheet processing software (Microsoft Excel for Windows).

RESULTS

There were no differences between the ACWR and the EW-ACWR calculations in relation to changes to workloads or RRI incidence proportions across the six ramp-up rates and the three random variation conditions under both MWP states (Supplementary Materiel Table 1). The EW-ACWR was, however, more sensitive than the traditional ACWR at responding to individual-level workload fluctuations (Figure 2). The differences between the 0.0% and 1.0%, and 1.0% and 2.5% random variation conditions did not considerably affect workloads or RRI incidence proportions (Supplementary Materiel Table 2). As such, we examined the interaction between the six ramp-up rates and the two most extreme random variation conditions, that is, 0.0% (perfect training adherence), and 2.5% (poor training adherence) under both MWP states.

<<<<< Insert Figure 2 about here >>>>>

Perfect training adherence (random variation 0.0%)

When the random variation was set to 0.0%, and the MWP state was set to on, the running population maintained the highest workloads relative to when the MWP was set to off (Figure 3). Similarly, higher ramp-up rates over the simulated timeframe also resulted in higher maximum workloads. Specifically, at a 5.0% and 30.0% ramp-up rate, the distance performed by the runners was 53.8km p/w and 62.5km p/w, respectively. Under the same set of conditions, the RRI incidence proportion was 0.0%. Conversely, with the MWP state set to off (i.e. runners could overshoot their MWP state), the RRI incidence proportion climbed from 4.2% to 30.1% across the six ramp-up rates.

<<<<< Insert Figure 3 about here >>>>>

The variability around workloads and RRI incidence proportions under the two different MWP states can be viewed in the dynamic ABM output plots (Figure 4 and 5). With the MWP set to on, the running population consistently increased their workload, and aware of the threshold over which they would sustain RRI, safely reached a performance ceiling (Figure 4). With the MWP state set to off, a given runner invariably surpassed their physical capacity and sustained a RRI (Figure 5).

<<<<< Insert Figure 4 and Figure 5 about here >>>>>

The workload across the six ramp-up rates remained relatively stable with the MWP state set to off (Figure 3). Accordingly, a 5.0% ramp-up rate resulted in runners spending a proportionately greater amount of time performing lower weekly running distances to the benefit of fewer RRIs (Figure 6). On the other hand, a 30.0% ramp-up rate reduced the length of time that runners spent at lower workloads, but equally resulted in a higher RRI incidence proportion. Across the population, MWP spikes stabilised with relatively longer periods of workload growth.

<<<<< Insert Figure 6 about here >>>>>

Poor training adherence (random variation 2.5%)

Introducing training error into the runners' ramp-up rates resulted in changes to both workloads and RRI incidence proportions (Figure 7). This condition simulated a scenario where runners were aware that a MWP state existed, but they could only estimate the value within a 2.5% random variation. At a 5.0% and 30.0% ramp-up rate, the mean distance performed by the runners was 29.9km p/w and 35.3km p/w, respectively. The RRI incidence proportion was higher across the six ramp-up rates relative to the 0.0% random variation condition.

<<<<< Insert Figure 7 about here >>>>>

With the random variation set to 2.5%, there was an initial upwards workload trajectory as runners climbed towards their MWP state (Figure 8). As a given runner approached and misjudged their MWP state due to training error, a higher RRI incidence proportion across the population brought the workload down over the 143 weeks.

<<<<< Insert Figure 8 about here >>>>>

DISCUSSION

The purpose of this study was to introduce computational modelling to sports injury research, and demonstrate the potential utility of ABM as a viable method for studying complex injury dynamics in future theoretical and practical applications. To achieve that purpose, an ABM was developed with the aim of simulating the dynamic relationship between the absolute weekly running distance and RRI, as well as the relative change to weekly running distance and RRI, through the manipulation of four athlete management tools (i.e. six ramp-up rates, three random variation conditions, two ACWR calculations, two MWP states). This was an important step in terms of complexity science and the systems thinking sports injury literature[1-4 6 7], particularly as no study has yet formally demonstrated the use of computational modelling in this context. Previous attempts to describe and/or apply the complex systems approach have resulted in the development of static frameworks or models that are not capable of simulating dynamic behaviours and emergent properties within complex sports systems, and/or understanding how systems change over time[14 15]. Aside from its practical offerings, this paper also reiterates the longstanding need for a paradigm shift toward 'dynamic simulations' and complex modelling[2 3]. Although the distance running ABM has effectively simulated the occurrence of sports injury in a population of runners, there remains a need to highlight what ABM can offer to the field of sports injury research more broadly. Therefore, the following discussion is structured around the main take-home messages, and subsequently outlines important considerations when aiming to use ABM in future research-based applications. The intention is to not only clarify the potential contribution of ABM, but also to inspire researchers and clinicians to continue to explore computational modelling and further develop applications in the sports injury context.

With regard to the presented simulation dynamics, a useful way of conceptualising the distance running ABM is to consider the rate at which the running population climbed towards a MWP state. Depending on the predefined ramp-up rate plus error condition, the goal of the agents was to run safely toward their maximum performance level. With the MWP state set to on, and the random variation condition set to 0.0% (i.e. perfect adherence to instruction), the running population appropriately identified their workload limits, and sustained the lowest number of RRIs. Conversely, increasing the random variation to 2.5% (i.e. poor adherence to instruction) whilst leaving all other conditions equal, adversely affected workloads and RRI incidence proportions over the course of the simulation. The maximum workload across the six ramp-up rates was comparatively lower when training error was higher because runners were misjudging the applied workload, and therefore sustaining RRI despite being aware of their MWP state.

Setting the MWP state to on and increasing the level of random variation in the model is representative of a real-world sports training situation. Distance runners, coaches, and qualified healthcare professionals are aware (or should be) that a MWP state exists, but knowing precisely where that upper individual limit is, and how to get there safely, is arguably one of the greatest challenges in prescribing future training loads. Despite the proven utility of workload calculation approaches such as the ACWR[40-46], the distance running ABM has demonstrated that building weekly running distances over time, even within reported sweet spot guidelines of up to 1.3[12 13 47], will eventually result in RRI as athletes reach their upper physical limits. This was supported by the simulation, as RRI was primarily generated by runners' surpassing their MWP state, and not due to extreme relative changes in the applied workload. Exceeding a physical capacity to tolerate workload is, however, not a new concept, and a 'ceiling effect of safety' has been described in the literature[51 52]. The findings of present study support the view that the calculation and prescription of athletic workloads should not be performed in isolation (i.e. with a single metric), and requires a comprehensive, individualised, and flexible approach[11 39 53].

Another insight offered by the simulation relates to the trade-off between a lower versus higher weekly ramp-up rate under the 2.5% random variation condition at different stages of recovery following RRI. Results showed that, in general, a higher ramp-up rate after returning from RRI resulted in greater maximum running distances over the course of the simulation as runners quickly returned to their pre-injury workload levels. For those athletes in the process of returning from RRI, this may be seen as positive. However, this relationship was offset by a higher RRI incidence proportion at the population-level. Specifically, in the early stages of recovery and prior to reaching their MWP state, runners experienced a greater margin of error when either miscalculating workload or not observing the recommended ramp-up rate. However, when runners' physical capacity to tolerate workload had been reached, any error to the applied running distance, irrespective of its magnitude, resulted in RRI. This 'hard ceiling' dynamic shows that it may be advantageous for runners who wish to maintain high distances over extended periods of time, to think long-term about their training, or perhaps even slightly refrain from wanting to regularly operate at their perceived level of peak performance. Whilst this simulation was modelled on understanding RRI occurrence in a 'general population', the implications of this dynamic indicated fragility at the extremities of performance for the more serious runner who might aim to participate in competitive events. Athletes, running coaches, and healthcare practitioners are reminded that although it is necessary to progressively and systematically increase external workloads over time, it is as equally important to continuously monitor and measure internal physiological and psychological responses to that load[39].

The conceptual basis and development of the distance running ABM was attributable to the complex systems approach[1]. For this reason, it is worthwhile to briefly outline the main distinction between computational modelling methods and more traditional analytical approaches for studying health-related multifactorial aetiologic mechanisms and complex systems phenomena. First and foremost, ABM is best used for exploring and understanding mechanisms and theories in complex systems that are unknown or contested[54]. Conversely, traditional statistical modelling approaches, such as regression analyses, are useful for testing *a priori* hypotheses and analysing already collected data. Although it is possible to analyse the results of a given ABM[55], it is not strictly a tool for analysing – it is a tool for understanding and generating hypotheses that can be empirically tested with statistical modelling. It is a way to dynamically simulate a range of complex systems characteristics that engender higher-level patterns that emerge from local, individual-level agent behaviours, thus enabling insight into how those patterns emerged with no 'top down' planning. As such, the application of ABM to a practical sports injury problem should be considered supplementary to epidemiologic inquiry, and requires a substantial epistemological and conceptual change in the way scientists and clinicians think about, and approach, the study of complex injury dynamics.

When studying complex systems phenomena with ABM, it is expected that a greater reliance is placed on theory relative to data[56]. That is, ABM cannot offer the same level of external quantitative credibility that traditional analytical approaches and statistical techniques can provide[57]. On the other hand, ABM enables the analyst to establish a balance between realism (i.e. face validity), generality (i.e. qualitative abstraction), and numerical precision (i.e. fineness of model specification). This balance can be achieved by triangulating different forms of evidence and using empirical data to parameterise models when assigning agent characteristics and environmental rules at baseline[58]. Notwithstanding the reported guidelines around the development, calibration, and validation of computational simulations[54 59], ABM is effectively an in-silico laboratory that can provide scientists and clinicians with a means of testing ideas and mechanisms of how their ideas may work in practice, albeit without undue financial, ethical, or logistical implications that are associated with real-world ‘pilot’ studies[55]. For instance, it is possible to situate agents within a social network and broader spatial context, duplicate initial baseline conditions, and subsequently change only one aspect of the model. Thus, a range of experimental scenarios can be trialled repeatedly, providing scientists with a powerful agent-based counterfactual simulation that evaluates the likely effect and impact of different health-related strategies and policies[55 56 60]. To be precise, ABM has been successfully used to emulate the randomised controlled trial for patients with diabetic retinopathy, allowing for the examination of hypothetical interventions targeting vision loss[61 62].

Given the flexibility and benefits of computational modelling, it is not surprising to find that ABM has been used to simulate cyclical, self-reinforcing feedback mechanisms among individual microunits, such as people, cells, and molecules, to identify emergent patterns of behaviour, such as disease transmission dynamics[63], wound healing processes[64 65], and adaptive immunity[66]. Indeed, one advantage of ABM is that the complex interactions between agents can (and ideally should) be explicitly modelled. Conversely, standard causal inference methods based on epidemiologic counterfactual reasoning are grounded in the Stable Unit Value Treatment Assumption, which explains that the effect of a given exposure and/or outcome of one individual is not affected by the exposure and/or outcome of others[35 55]. However, depending on the context and scope of analysis, this assumption is often violated when there is interference among units leading to biased causal effects[67] (e.g. athletes are capable of influencing the behaviours of their peers). A final consideration when using certain regression-based analyses (e.g. logistic regression, survival analyses), and something to which ABM can potentially circumnavigate, is the events-per-variable requirement[68 69]. In the absence of large scale data sets, other methods such as ABM might prove useful for simulating hundreds or thousands of agents, each of whom can be assigned personal characteristics and behaviours. Simulating a considerably large population might lead to a sufficiently large number of injurious events per explanatory variable modelled, affording insight into possible mechanisms that generate certain outcomes. Although beyond the scope of this paper, a more complete list of the advantages and disadvantages of ABM should be provided and contextualised within the sports science field.

On the whole, and in terms of sports injury research, ABM has the potential to test supposed aetiologic mechanisms, generate new causal hypotheses, investigate the extent to which different factors and their interactions influence the onset of injury, or examine the impact of new hypothetical injury prevention strategies within considerably large synthetic athlete populations. Irrespective of why ABM is used, its development has to be carefully planned, debated, and scrutinised over a series of iterative stages that starts with a verified working model, similar to the distance running ABM presented. There is now a need for future computational modelling applications to explore how ABM can be used to simulate more advanced complex systems characteristics in relation to sports injury aetiology and prevention.

Limitations and research-based considerations

This proof-of-concept study is not without limitation. First and foremost, the intention of the ABM was not to offer original data nor provide practical knowledge about how to safely increase workloads for running performance optimisation. For example, runners' personal characteristics such as BMI were not explanatory in the sense of impacting on the results, and this should motivate future computational applications to either build on the distance running ABM directly (Supplementary NetLogo Code provided), or draw upon its premise to guide the development of dynamic simulations in other sports domains. For that reason, the model was not instantiated with data, and the selection of runners' personal characteristics, as well as numerical values pertaining to workload and RRI risk, were based on subject matter expertise (AH and RN), contemporary RRI causal theory[9 10], and evidence around the ACWR[12 13]. We contend that this provides a practical position from which to further explore computational modelling.

Another limitation relates to the different classes of agent-based computational simulations that can be developed. For example, the distance running ABM is more representative of an engineering-theoretic model, as runners responded to their environment (i.e. direction from the coach) but operated independently from one another. Conversely, a typical ABM in the social sciences aims to understand how the mechanism of interaction between boundedly rational agents leads to the emergence of global patterns and collective behaviour. Extending the current model to include local level interactions between agents and factors could be a means of providing further insight into the role of specific mechanisms that drive behaviour and injury incidence, as well as recovery among the running population. Although there are many different ways to advance the distance running ABM, it is essential that extensions are biologically plausible, theoretically reasonable, and numerically precise where relevant[57].

In terms of research-based considerations, ABM is arguably more amenable to the study of infectious disease. This is because agent interactions are predicated on known contagion dynamics, and can be bolstered with historic epidemic data. Conversely, developing models for noncommunicable health-related conditions requires a degree of ingenuity, and this proved challenging in the current study. Developing ABM requires the expert use of an object-oriented programming language (e.g. Java, Python, C++), and the wide range of available computer-based software packages could be disconcerting. Overcoming these hurdles and transitioning to computational modelling requires a multidisciplinary team comprising epidemiologists, biostatisticians, and persons with proven level of expertise in coding and programming.

CONCLUSIONS

It has long been argued that the adoption of a complex systems approach in sports injury research and practice will provide a paradigm shift in how injury is understood and prevented. However, beyond theoretical description and static models of complexity, little progress has been made in formalising this approach in a way that is practical and meaningful to sports injury scientists and clinicians. In short, a true complex systems approach has not yet been applied. However, in demonstrating for the first time the use of ABM to simulate RRI, this study provides a compelling case for the use of complex systems modelling methods in the sports injury context moving forwards. The distance running ABM was underpinned by contemporary RRI causal theory, and incorporates the ACWR, an approach that calculates relative changes in workload. The findings indicate that the model can simulate fundamental running workload and RRI dynamics in a simulated population. Specifically, when runners were aware of their upper workload limits, and perfectly adhered to training-related instruction, the running population maintained the highest level of performance whilst sustaining the lowest number of RRIs. Conversely, under the same set of conditions, poor adherence to training-related instruction adversely affected workloads and RRI incidence proportions over the course of the simulation. The model also served to confirm that the calculation and prescription of athletic workloads should not rely upon the use of single metric, and requires a comprehensive, personalised, and adaptable approach. This is especially true for runners who are operating close to their physiological potential. The implications of this study, however, extend beyond the presented model. In particular, scientists and clinicians interested in the philosophy of complex systems and their associated emergent dynamics should start to explore what ABM can offer to a specific sports injury problem of interest. Alongside the continuing use of traditional epidemiological and clinical research-based applications, the use of ABM should now occupy a methodological and analytical space in sports medicine research.

What are the findings?

- There was a need to introduce complex systems computational modelling to the field of sports injury research. The use of methods such as ABM are complementary alternatives to routine analytic approaches which can be limited by traditional statistical assumptions.
- This proof-of-concept study applies Agent-Based Modelling (ABM) in the context of running-related injury (RRI). The findings show that the distance running ABM can simulate fundamental running workload and RRI dynamics.
- The simulation demonstrated that although it is important to maintain workloads within the ACWR sweet spot, there will come a point whereby an individual ceiling of safety is surpassed resulting in RRI. This dynamic is exacerbated with greater workload error.

How might it impact on clinical practice in the near future?

- Computational modelling methods such as ABM are primarily used to understand how local-level behaviours and interactions among individual 'agents' (e.g. molecules, cells, people) leads to the emergence of complex systems patterns (e.g. sports injury development).
- Sports injury scientists and clinicians should familiarise themselves with ABM to determine new ways of using this theoretically-driven method so that it can be effectively applied to a specific problem of interest.
- In the absence of large-scale data, scientists and clinicians interested in the aetiology and prevention of sports injury should consider ABM as an alternative and complementary method to traditional epidemiological and clinical research-based applications.

ACKNOWLEDGEMENTS

We would like to thank Professor Rod McClure for his insightful and constructive feedback during the early stages of this research. We also thank Mr Tony Carden whose knowledge on the origins and history of complexity theory was most valuable. Dr Simon Feros and Dr Scott Mclean provided useful comments pertaining to the flow and readability of the manuscript.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHOR CONTRIBUTIONS

Adam Hulme was responsible for the idea, concept, ABM development, methods, results interpretation, and write-up. Jason Thompson was primarily responsible for developing the ABM, contributed to the methods write-up, and had editorial input into the manuscript. Rasmus Nielsen, Gemma Read, and Paul Salmon had editorial input into the manuscript, and contributed to the write-up.

References

1. Bittencourt NFN, Meeuwisse, W.H, Mendonça, L.D, Nettel-Aguirre, A, Ocarino, J.M, Fonsesca, S.T. Complex systems approach for sports injuries. Moving from risk factor identification to injury pattern recognition - a narrative review and new concept. *Br J Sports Med* 2016;0:1-7.
2. Quatman CE, Quatman CC, Hewett TE. Prediction and prevention of musculoskeletal injury: a paradigm shift in methodology. *Br J Sports Med* 2009;43(14):1100-07. doi: 10.1136/bjsm.2009.065482
3. Mendiguchia J, Alentorn-Geli E, Brughelli M. Hamstring strain injuries: are we heading in the right direction? *Br J Sports Med* 2011 doi: 10.1136/bjsm.2010.081695
4. Hulme A, Finch, C.F. From monocausality to systems thinking: A complementary and alternative conceptual approach for better understanding the development and prevention of sports injury. *Injury Epidemiology* 2015;2(1):1-12. doi: 10.1186/s40621-015-0064-1
5. Bertalanffy L. General system theory: Foundations, development, application. New York: Brazilier 1969.
6. Bekker S, Clark, A.M. Bringing complexity to sports injury prevention research: from simplification to explanation. *Br J Sports Med* 2016 doi: doi:10.1136/bjsports-2016-096457
7. Mooney M, Charlton, P.C, Soltanzadeh, S, Drew, M.K. Who 'owns' the injury or illness? Who 'owns' performance? Applying systems thinking to integrate health and performance in elite sport *Br J Sports Med* 2017 doi: 10.1136/bjsports-2016-096649
8. Hulme A, Finch, C.F. The epistemic basis of distance running injury research: A historical perspective. *Journal of Sport and Health Science* 2016;5(2):172-75. doi: doi:10.1016/j.jshs.2016.01.023
9. Malisoux L, Nielsen, R.O, Urhausen, A, Theisen, D. A step towards understanding the mechanisms of running-related injuries. *J Sci Med Sport* 2015;18(5):523-28.
10. Bertelsen ML, Hulme, A, Petersen, J, Brund, R.K, Sørensen, H, Finch, C.F, Parner, E.T, Nielsen, R.O. . A framework for the etiology of running-related injuries. *Scand J Med Sci Sports* 2017 doi: doi: 10.1111/sms.12883
11. Nielsen RO, Bertelsen, M.L, Møller, M, Hulme, A, Windt, J, Verhagen, E.A, Mansournia, M.A, Casals, M, Parner, E.T. Training load and structure-specific load: applications for sport injury causality and data analyses. *Br J Sports Med* 2017 doi: 10.1136/bjsports-2017-097838
12. Gabbett TJ. The training-injury prevention paradox: should athletes be training smarter and harder? *Br J Sports Med* 2016;50(5):273-80.
13. Gabbett TJH, B.T. Blanch, P, Whiteley, R. High training workloads alone do not cause sports injuries: how you get there is the real issue. *Br J Sports Med* 2016;50(8):444-45.
14. Hulme A, Salmon, P.M, Nielsen, R.O, Read, G.L.M, Finch, C.F. Closing Pandora's Box: Adapting a systems ergonomics methodology for better understanding the ecological complexity underpinning the development and prevention of running-related injury. *Theoretical Issues in Ergonomics Science* 2017;18(4):338-59. doi: 10.1080/1463922X.2016.1274455
15. Hulme A, Salmon, P.M, Nielsen, R.O, Read, G.J.M, Finch, C.F. From control to causation: Validating a 'complex systems model' of running-related injury development and prevention. *Appl Ergon* 2017 doi: 10.1016/j.apergo.2017.07.005
16. Leveson NG. A new accident model for enginnering safer systems. *Safety Science* 2004;42(4):237-70.
17. Steves R, Hootman JM. Evidence-Based Medicine: what is it and how does it apply to athletic training? *J Athl Train* 2004;39(1):83-87.

18. Iles R, Davidson M. Evidence based practice: a survey of physiotherapists' current practice. *Physiother Res Int* 2006;11(2):93-103. [published Online First: 2006/07/01]
19. Manske RC, Lehecka BJ. Evidence-Based Medicine/practice in sports physical therapy. *International Journal of Sports Physical Therapy* 2012;7(5):461-73.
20. Auchincloss AH, Diez Roux, A.V. A new tool for epidemiology: the usefulness of dynamic-agent models in understanding place effects on health. *Am J Epidemiol* 2008;168(1):1-8.
21. Moore D, Dray, A, Green, R, Hudson, S.L, Jenkinson, R, Siokou, C, Perez, P, Bammer, G, Maher, L, Dietze, P. Extending drug ethno-epidemiology using agent-based modelling. *Addiction* 2009;104(12):1991-97.
22. Siettos C, Anastassopoulou C, Russo L, et al. Modeling the 2014 Ebola Virus Epidemic - Agent-Based Simulations, Temporal Analysis and Future Predictions for Liberia and Sierra Leone. *PLoS currents* 2015;7 doi: 10.1371/currents.outbreaks.8d5984114855fc425e699e1a18cdc6c9 [published Online First: 2015/06/13]
23. Venkatramanan S, Lewis B, Chen J, et al. Using data-driven agent-based models for forecasting emerging infectious diseases. *Epidemics* 2017 doi: http://dx.doi.org/10.1016/j.epidem.2017.02.010
24. Bonabeau E. Agent-based modeling: Methods and techniques for simulating human systems. *Proc Natl Acad Sci U S A* 2002;99(3):7280-87.
25. Macal CM, North, M.J. Tutorial on agent-based modelling and simulation. *Journal of Simulation* 2010;4(3):151-62.
26. Epstein JM. Modelling to contain pandemics. *Nature* 2009;460(7256):460-687.
27. Parker J, Epstein JM. A distributed platform for global-scale agent-based models of disease transmission. *ACM transactions on modeling and computer simulation : a publication of the Association for Computing Machinery* 2011;22(1):2. doi: 10.1145/2043635.2043637 [published Online First: 2011/12/01]
28. Luke DA, Stamatakis, K.A. Systems science methods in public health. *Annu Rev Public Health* 2012;33:357-76.
29. Perez L, Dragicevic S. An agent-based approach for modeling dynamics of contagious disease spread. *International journal of health geographics* 2009;8:50. doi: 10.1186/1476-072x-8-50 [published Online First: 2009/08/07]
30. Nianogo RA, Arah, O.A. Agent-based modeling of noncommunicable diseases: a systematic review. *Am J Public Health* 2015;105(3):e20-31.
31. Yang Y, Diez Roux, A.V, Auchincloss, A.H, Rodriguez, D.A, Brown, D.G. A spatial agent-based model for the simulation of adults' daily walking within a city. *American Journal of Preventative Medicine* 2011;40(3):353-61.
32. Auchincloss AH, Riolo, R.L, Brown, D.G, Cook, J, Roux Diez, A.V. An agent-based model of income inequalities in diet in the context of residential segregation. *American Journal of Preventative Medicine* 2013;40(3):303-11.
33. Zhang D, Giabbanelli, P.J, Arah, O.A, Zimmerman, F.J. Impact of different policies on unhealthy dietary behaviors in an urban adult population: An agent-based simulation model. *Am J Public Health* 2014;104(7):1217-22.
34. Eddy DM, Schlessinger L, Kahn R. Clinical outcomes and cost-effectiveness of strategies for managing people at high risk for diabetes. *Ann Intern Med* 2005;143(4):251-64. doi: 10.7326/0003-4819-143-4-200508160-00006

35. Marshall BM. Agent-based modelling. Systems Science and Population Health: Oxford University Press, New York 2017:87-98.
36. Wilensky U. NetLogo 2016 [Available from: https://ccl.northwestern.edu/netlogo/index.shtml accessed 21 August 2017.
37. Netlogo. 2017 [Available from: https://ccl.northwestern.edu/netlogo/ accessed 22 August 2017.
38. Schwellnus M, Soligard T, Alonso J-M, et al. How much is too much? (Part 2) International Olympic Committee consensus statement on load in sport and risk of illness. *Br J Sports Med* 2016;50(17):1043-52. doi: 10.1136/bjsports-2016-096572
39. Soligard T, Schwellnus M, Alonso J-M, et al. How much is too much? (Part 1) International Olympic Committee consensus statement on load in sport and risk of injury. *Br J Sports Med* 2016;50(17):1030-41. doi: 10.1136/bjsports-2016-096581
40. Hulin BT, Gabbett TJ, Blanch P, et al. Spikes in acute workload are associated with increased injury risk in elite cricket fast bowlers. *Br J Sports Med* 2013 doi: 10.1136/bjsports-2013-092524
41. Hulin BT, Gabbett TJ, Lawson DW, et al. The acute:chronic workload ratio predicts injury: high chronic workload may decrease injury risk in elite rugby league players. *Br J Sports Med* 2015 doi: 10.1136/bjsports-2015-094817
42. Hulin BT, Gabbett TJ, Caputi P, et al. Low chronic workload and the acute:chronic workload ratio are more predictive of injury than between-match recovery time: a two-season prospective cohort study in elite rugby league players. *Br J Sports Med* 2016;50(16):1008-12. doi: 10.1136/bjsports-2015-095364
43. Carey DL, Blanch P, Ong K-L, et al. Training loads and injury risk in Australian football—differing acute: chronic workload ratios influence match injury risk. *Br J Sports Med* 2016 doi: 10.1136/bjsports-2016-096309
44. Malone S, Roe M, Doran DA, et al. High chronic training loads and exposure to bouts of maximal velocity running reduce injury risk in elite Gaelic football. *J Sci Med Sport* 2017;20(3):250-54. doi: http://dx.doi.org/10.1016/j.jsams.2016.08.005
45. Malone S, Roe M, Doran DA, et al. Protection Against Spikes in Workload With Aerobic Fitness and Playing Experience: The Role of the Acute:Chronic Workload Ratio on Injury Risk in Elite Gaelic Football. *International journal of sports physiology and performance* 2017;12(3):393-401. doi: 10.1123/ijspp.2016-0090 [published Online First: 2016/07/12]
46. Bowen L, Gross AS, Gimpel M, et al. Accumulated workloads and the acute:chronic workload ratio relate to injury risk in elite youth football players. *Br J Sports Med* 2017;51(5):452-59. doi: 10.1136/bjsports-2015-095820
47. Windt J, Gabbett, T.J. How do training and competition workloads relate to injury? The workload—injury aetiology model. *Br J Sports Med* 2016;51(5):428-35.
48. Menaspà P. Are rolling averages a good way to assess training load for injury prevention? *Br J Sports Med* 2016 doi: 10.1136/bjsports-2016-096131
49. Murray NB, Gabbett TJ, Townshend AD, et al. Calculating acute:chronic workload ratios using exponentially weighted moving averages provides a more sensitive indicator of injury likelihood than rolling averages. *Br J Sports Med* 2017;51(9):749-54. doi: 10.1136/bjsports-2016-097152
50. Williams S, West S, Cross MJ, et al. Better way to determine the acute:chronic workload ratio? *Br J Sports Med* 2017;51(3):209-10. doi: 10.1136/bjsports-2016-096589
51. Drew MK, Cook J, Finch CF. Sports-related workload and injury risk: simply knowing the risks will not prevent injuries. *Br J Sports Med* 2016 doi: 10.1136/bjsports-2015-095871 [published Online First: 2016/05/12]

52. Stares J, Dawson B, Peeling P, et al. Identifying high risk loading conditions for in-season injury in elite Australian football players. *J Sci Med Sport* 2018;21(1):46-51. doi: 10.1016/j.jsams.2017.05.012 [published Online First: 2017/06/12]
53. Drew MK, Blanch P, Purdam C, et al. Yes, rolling averages are a good way to assess training load for injury prevention. Is there a better way? Probably, but we have not seen the evidence. *Br J Sports Med* 2017;51(7):618-19. doi: 10.1136/bjsports-2016-096609
54. Rand W, Rust RT. Agent-based modeling in marketing: Guidelines for rigor. *International Journal of Research in Marketing* 2011;28(3):181-93. doi: https://doi.org/10.1016/j.ijresmar.2011.04.002
55. Marshall BDL, Galea, S. Formalizing the role of agent-based modeling in causal inference and epidemiology. *Am J Epidemiol* 2014;181(2):92-99.
56. Hernán MA. Invited commentary: Agent-based models for causal inference - reweighting data and theory in epidemiology. *Am J Epidemiol* 2015;181(2):103-05.
57. Ip EH, Rahmandad, H, Shoham, D.A, Hammond, R, Huang, T.K, Wang, Y, Mabry, P.L. Reconciling statistical and systems science approaches to public health. *Health Educ Behav* 2013;40(1):123-31.
58. Berk R. How you can tell if the simulations in computational criminology are any good. *Journal of Experimental Criminology* 2008;4(3):289. doi: 10.1007/s11292-008-9053-5
59. Caro JJ, Briggs AH, Siebert U, et al. Modeling good research practices—overview: A report of the ISPOR-SMDM modeling good research practices task force-1. *Value Health* 2012;15(6):796-803. doi: https://doi.org/10.1016/j.jval.2012.06.012
60. El-Sayed AM, Scarborough, P, Seeman, L, Galea, S. Social network analysis and agent-based modeling in social epidemiology. *Epidemiologic Perspectives and Innovations* 2012;9(1):1-9.
61. Day TE, Ravi N, Xian H, et al. An Agent-Based Modeling Template for a Cohort of Veterans with Diabetic Retinopathy. *PLoS ONE* 2013;8(6):e66812. doi: 10.1371/journal.pone.0066812
62. Day TE, Ravi N, Xian H, et al. Sensitivity of diabetic retinopathy associated vision loss to screening interval in an agent-based/discrete event simulation model. *Comput Biol Med* 2014;47:7-12. doi: 10.1016/j.compbiomed.2014.01.007 [published Online First: 2014/02/11]
63. Marshall BDL, Paczkowski, M.N, Seemann, L, Tempalski B, Pouget, R.R, Galea, S, Friedman, S.R. A complex systems approach to evaluate HIV prevention in metropolitan areas: Preliminary implications for combination intervention strategies. *PLoS ONE* 2012;7(9) doi: doi: 10.1371/journal.pone.0044833
64. Mi Q, Rivière, B, Clermont, G, Steed, D.L, Vodovotz, Y. Agent-Based model of inflammation and wound healing: insights into diabetic foot ulcer pathology and the role of transforming growth factor-beta1. *Wound Repair Regen* 2007;15(5):671-82.
65. Gopalakrishnan V, Kim M, An G. Using an Agent-Based Model to Examine the Role of Dynamic Bacterial Virulence Potential in the Pathogenesis of Surgical Site Infection. *Adv Wound Care* 2013;2(9):510-26. doi: 10.1089/wound.2012.0400
66. Folcik VA, An GC, Orosz CG. The Basic Immune Simulator: an agent-based model to study the interactions between innate and adaptive immunity. *Theor Biol Med Model* 2007;4:39. doi: 10.1186/1742-4682-4-39 [published Online First: 2007/09/29]
67. Oakes JM, Naimi AI. Mediation, interaction, interference for social epidemiology. *Int J Epidemiol* 2016;45(6):1912-14. doi: 10.1093/ije/dyw279
68. Hansen SN, Andersen PK, Parner ET. Events per variable for risk differences and relative risks using pseudo-observations. *Lifetime Data Anal* 2014;20(4):584-98. doi: 10.1007/s10985-013-9290-4 [published Online First: 2014/01/15]

69. Peduzzi P, Concato J, Feinstein AR, et al. Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. *J Clin Epidemiol* 1995;48(12):1503-10. [published Online First: 1995/12/01]

FIGURES AND LEGENDS

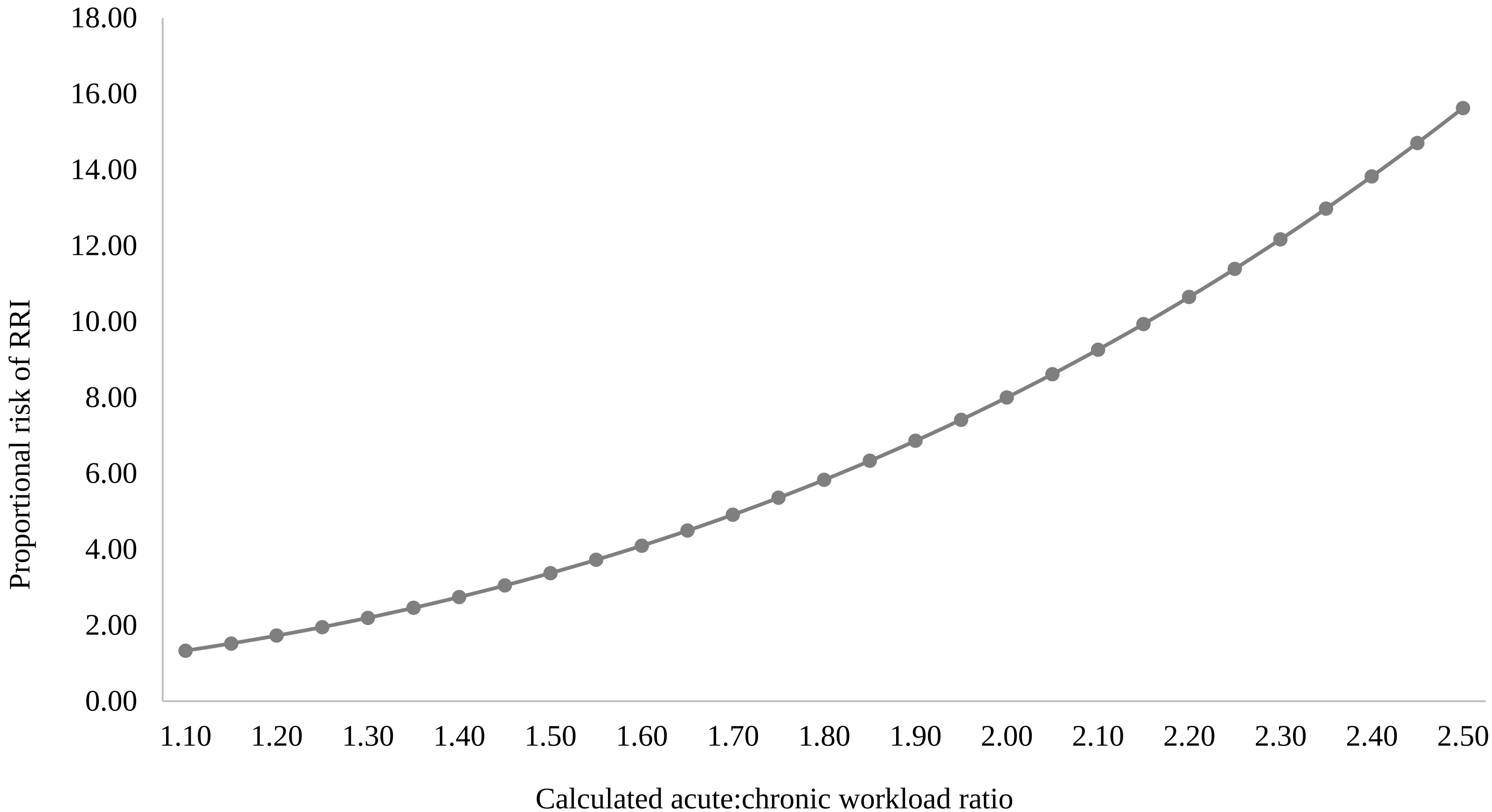

Figure 1: Proportional increase in RRI risk with an increasing ACWR. Higher ratios resulted in an exponentially increasing risk of RRI.

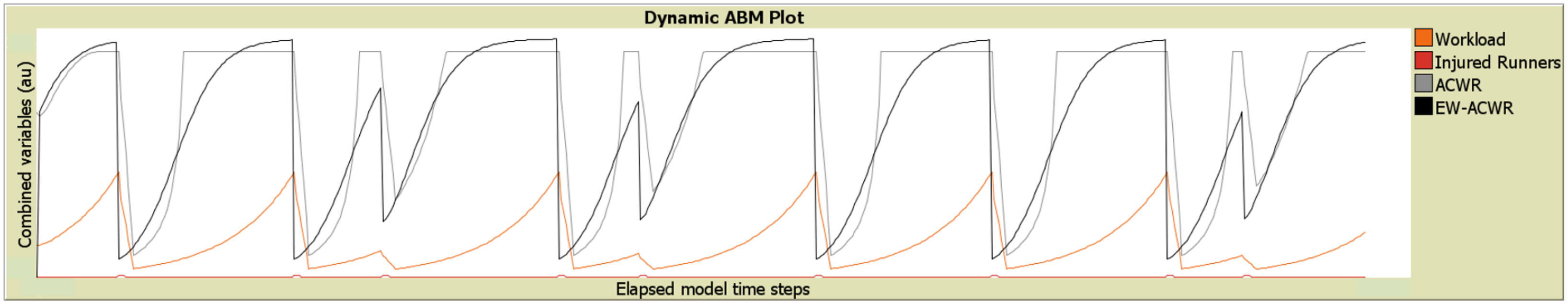

Figure 2: Dynamic plot visualising the relationship between workload and RRI over 80 weeks for a single runner. Ramp-up rate set to 30.0%, random variation set to 0.0%, MWP set to off. The black, grey, orange, and red lines represent the EW-ACWR, the traditional ACWR, workload, and RRI incidence proportion, respectively. Unlike the ACWR which exhibited smoothing behaviour due to the calculation of rolling averages, the EW-ACWR closely tracked workloads, and rapidly dropped off when RRI was sustained. Note that at three separate time points, the absolute workload fell short of the runner's MWP state, indicating that RRI was sustained by chance before the runner could reach their absolute upper workload limit. The use of a single runner and 80 weeks (570 elapsed model time steps) is for illustrative purposes.

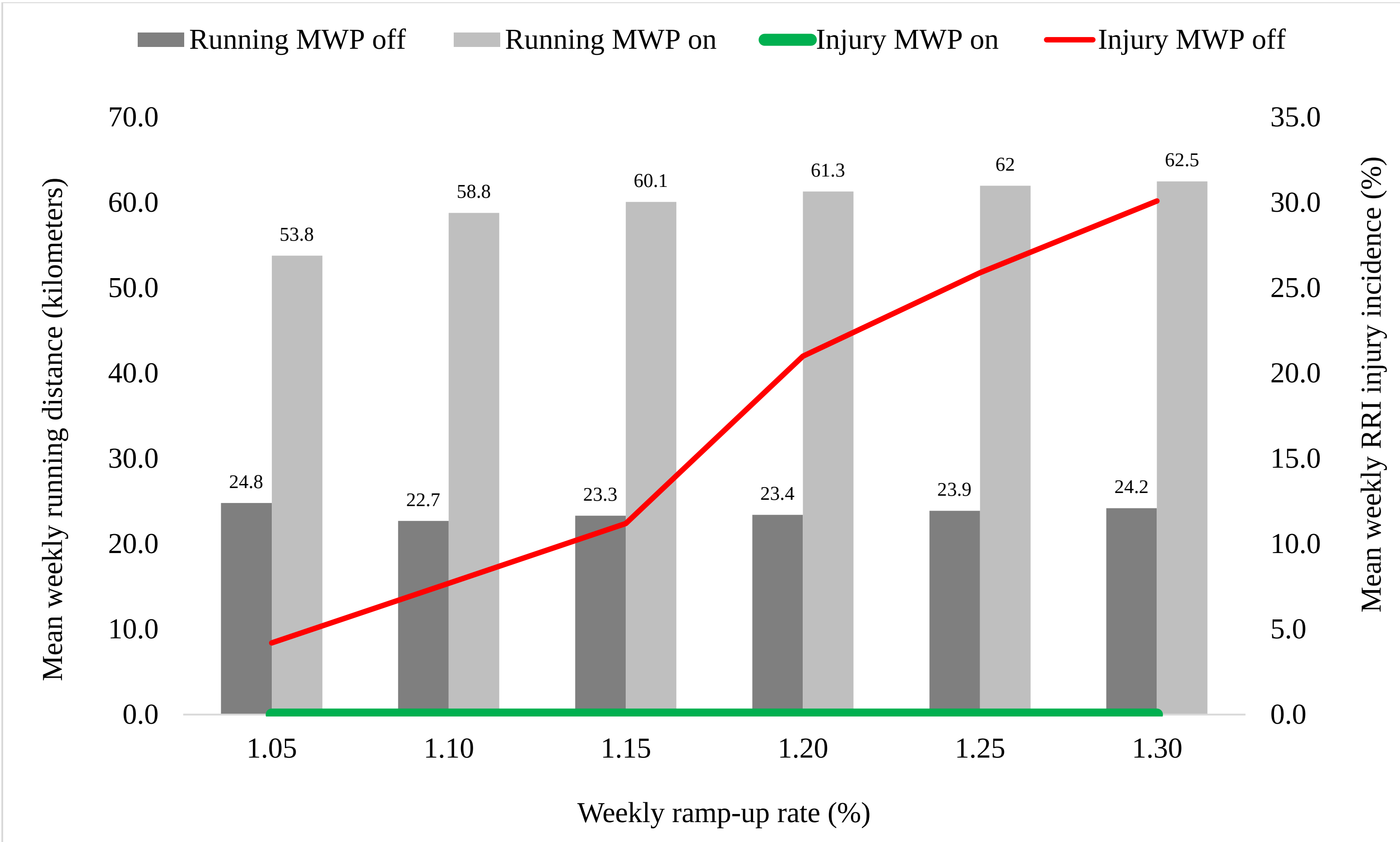

Figure 3: A comparison of workloads and RRI incidence proportions under the two MWP states across six ramp-up rate conditions. Random variation was set to 0.0% over 143 weeks.

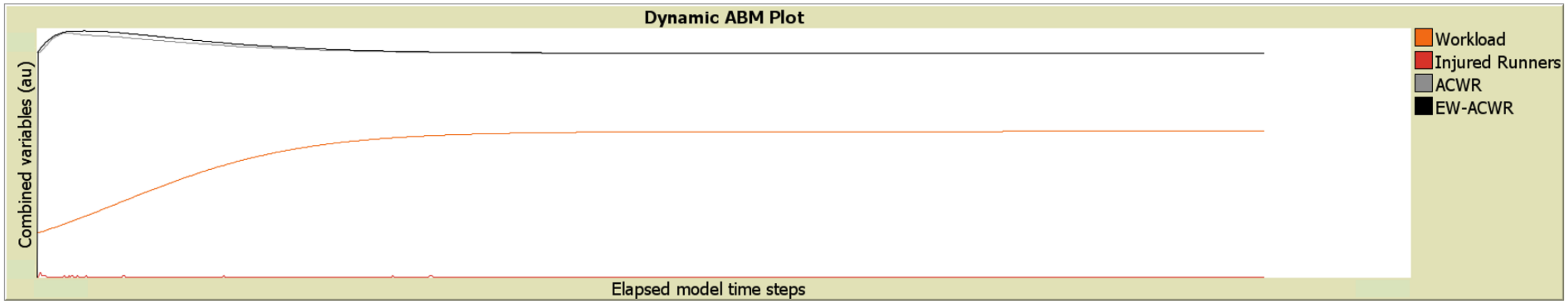

Figure 4: Dynamic plot visualising the relationship between workload and RRI over 143 weeks for 1000 runners.

Ramp-up rate set to 10.0%, random variation set to 0.0%, and MWP set to on. The black, grey, orange, and red lines represent the EW-ACWR, the traditional ACWR, workload, and RRI incidence proportion, respectively. As runners neared a MWP state, the upper bound of the ACWR inwardly collapsed to 1.0.

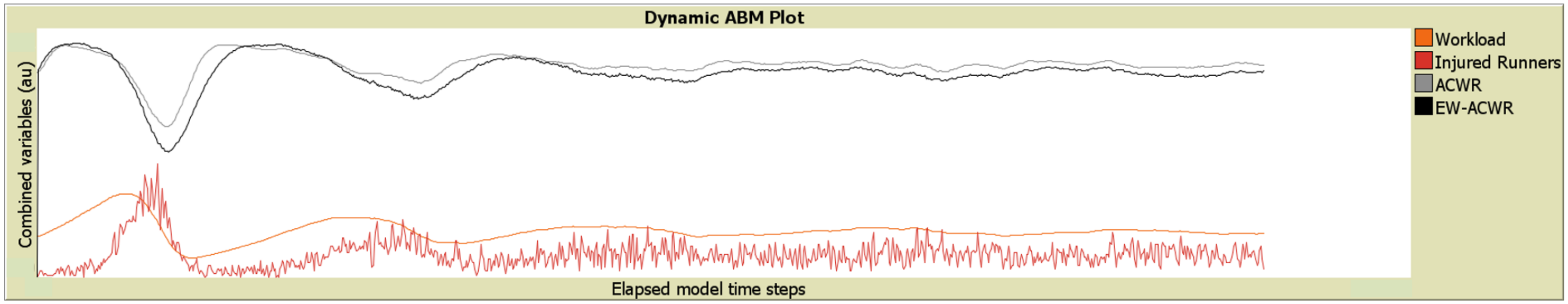

Figure 5: Dynamic plot visualising the relationship between workload and RRI over 143 weeks for 1000 runners. Ramp-up rate set to 10.0%, random variation set to 0.0%, and MWP set to off. The black, grey, orange, and red lines represent the EW-ACWR, the traditional ACWR, workload and RRI incidence proportion, respectively. The initial oscillatory pattern stabilised as runners sustained RRI at different times due to a randomly distributed MWP state.

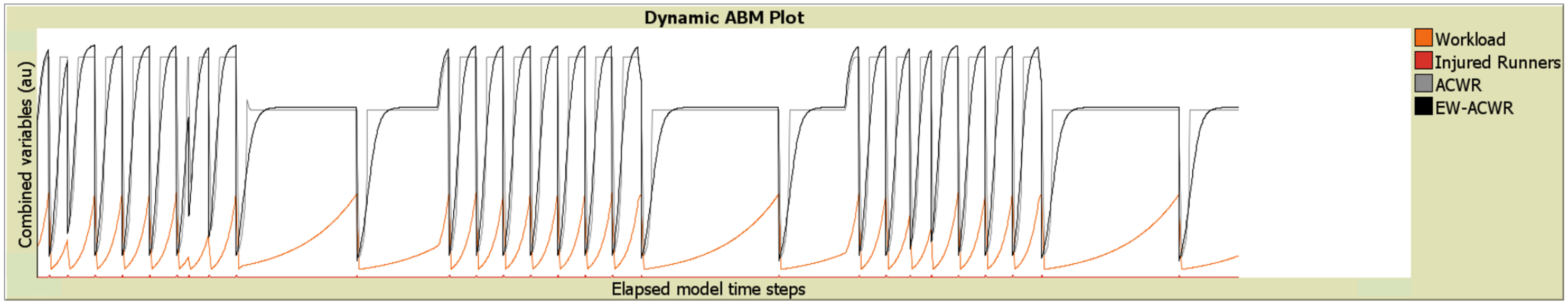

Figure 6: Dynamic plot visualising the relationship between workload and RRI over 428 weeks for a single runner. Random variation set to 0.0%, and MWP state set to off. The black, grey, orange, and red lines represent the EW-ACWR, the traditional ACWR, workload, and RRI incidence proportion, respectively. The output plot visualises the changes to workload and the RRI incidence proportion when the ramp-up rate is switched, on the fly, between 30.0% and 5.0% every ~71 weeks. Note the greater workload oscillation at a 30.0% relative to a 5.0% ramp-up rate. The use of a single runner and 428 weeks (3000 elapsed model time steps) is for illustrative purposes.

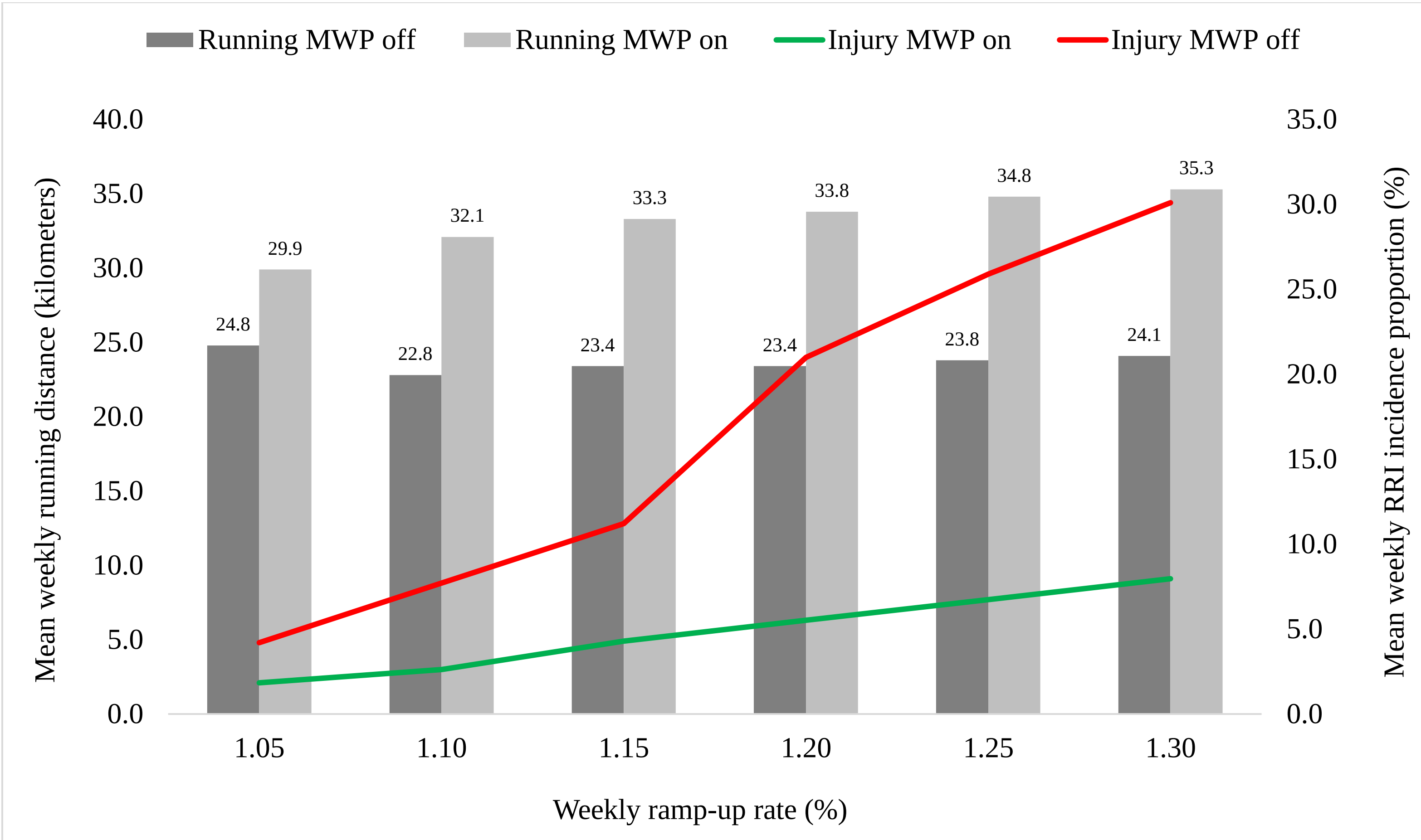

Figure 7: A comparison of workloads and RRI incidence proportions under the two MWP states across six ramp-up rate conditions. Random variation was set to 2.5% over 143 weeks.

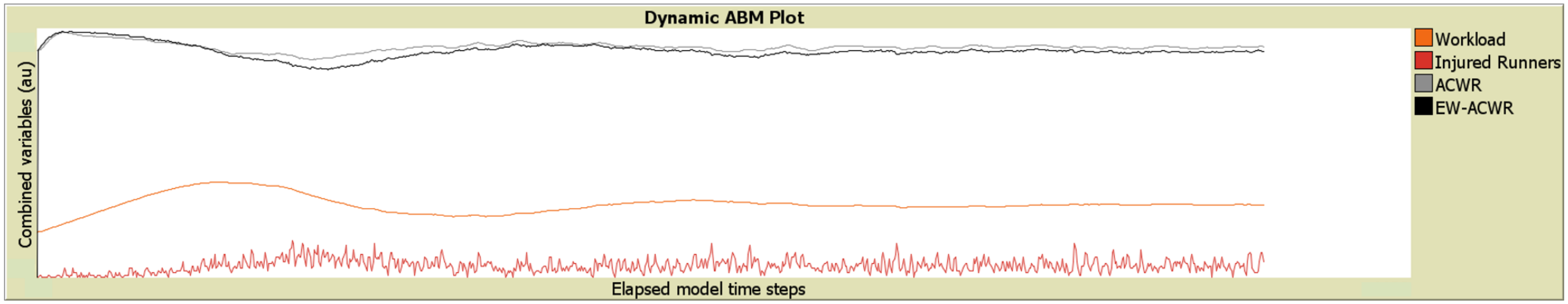

Figure 8: The relationship between workload and RRI over 143 weeks for 1000 runners. Ramp-up rate set to 10.0%, random variation set to 2.5%, and MWP set to on. The black, grey, orange, and red lines represent the EW-ACWR, the traditional ACWR, workload, and RRI incidence proportion, respectively.